# Never Underestimate Your Opponent: Hindsight Bias Causes Overplacement and Overentry into Competition[*]

David Danz[†]

October 27, 2020.[‡]

**Abstract**

This study investigates a source of comparative overconfidence, or overplacement, which occurs when people overestimate themselves relative to others. We present a simple application of information projection (Madarász, 2012) to show that hindsight bias can lead to overplacement and excessive willingness to compete. We run an experiment in which subjects choose between a competitive tournament and piece-rate compensation after observing some of their competitors' past performance. We exogenously manipulate whether subjects have ex post information about their competitors' past tasks (hindsight) or not (no hindsight). We find that hindsight bias generates overplacement and increases subjects' valuation of tournament participation by 19%. In line with theory, the additional tournament entry in the hindsight setting is driven by low-performing participants who should not have entered the tournament.

*Keywords*: Overconfidence; tournament; hindsight bias; curse of knowledge; information projection; beliefs; entrepreneurship.

[*]I am grateful to Valeria Burdea, Dietmar Fehr, Jana Friedrichsen, Holger Gerhardt, Osea Giuntella, Luise Görges, Rustamdjan Hakimov, David Huffman, Steffen Huck, Radosveta Ivanova-Stenzel, Tobias König, Dorothea Kübler, Sera Linardi, Kristóf Madarász, Lydia Mechtenberg, Muriel Niederle, Veronica Rattini, Jennifer Rontganger, Ronald Rensink, Julia Schmid, Justin Valasek, Roel van Veldhuizen, Lise Vesterlund, Stephanie Wang, Georg Weizsäcker, Axel Werwatz, anonymous reviewers, and seminar audiences at HWR Berlin, WZB-TU Colloquium, BBE Workshop 2013, CRC 649 Workshop 2013, IMEBESS 2014, CIRPÉE-TU-WZB Workshop 2014, and Stanford SITE 2015 for valuable comments and discussions. Financial support from the Deutsche Forschungsgemeinschaft (DFG) through SFB 649 "Economic Risk" is gratefully acknowledged.

[†]University of Pittsburgh, Department of Economics, 230 S. Bouquet St., 4922 Posvar Hall, Pittsburgh, PA 15260, and WZB Berlin. Email: danz@pitt.edu.

[‡]An earlier version of this paper appeared under the title "The curse of knowledge increases self-selection into competition: Experimental evidence" (Danz, 2014); First version: Danz (2013).

## 1 Introduction

Overconfidence is a widespread phenomenon and is found among managers (Huffman et al., 2019), CEOs (Malmendier and Tate, 2005, 2015), professional traders (Glaser et al., 2005), and entrepreneurs (Cooper et al., 1988). It has been linked to excessive trading (Odean, 1998; Barber and Odean, 2001), speculative bubbles (Scheinkman and Xiong, 2003), bargaining impasse (Neale and Bazerman, 1985), poor acquisitions (Hayward and Hambrick, 1997; Malmendier and Tate, 2008), entrepreneurial failure (Wu and Knott, 2006; Koellinger et al., 2007), and war (Howard, 1984; Johnson, 2004).

In this paper, we seek to better understand the conditions that lead to comparative overconfidence, or overplacement, which occurs when people overestimate themselves relative to others. In particular, we consider a factor that has not been tested directly as a cause of overplacement: hindsight bias. We investigate whether hindsight bias leads to comparative overconfidence in beliefs and to overconfident behavior in the form of overentry into competitive tournaments. As a mechanism, hindsight bias makes interesting predictions about when overconfidence will be particularly pronounced, namely situations characterized by asymmetric information. Paradoxically, more information can worsen overplacement through hindsight bias. Hindsight bias also predicts that novice actors who have little experience with a task are particularly prone to overconfidence.

Hindsight bias might be an important and overlooked factor for overplacement in real-world settings. First, as "[o]ne of the most widely studied biases in the judgment literature" (Rabin, 1998), hindsight bias is hard to overcome with debiasing techniques (Guilbault et al., 2004). It is found in, among others, bankers (Biais and Weber, 2009), physicians (Arkes et al., 1981; Dawson et al., 1988; Arkes, 2013), judges (Anderson et al., 1997; Harley, 2007), and entrepreneurs (Cassar and Craig, 2009). Second, previous research in the lab on overplacement and overentry has been conducted in settings with symmetric information about subjects' tasks, where hindsight bias cannot occur. In the real world, informational asymmetries emerge naturally over time and performance evaluations are thus likely to be affected by hindsight bias.

Our predictions are based on a simple application of Madarász's (2012) model of *information projection*. Information projection unifies a number of well-documented biases that occur under informational asymmetries: hindsight bias (Fischhoff, 1975; Fischhoff and Beyth, 1975), curse of knowledge (Camerer et al., 1989), and illusion of transparency (Gilovich et al., 1998). These biases have in common that they describe our inability to put ourselves in the shoes of less-informed others (or our past self). With hindsight bias, we think that we "knew it all along" and overestimate the prior predictability of historic events (Fischhoff, 1975; Fischhoff and Beyth, 1975), elections (Blank et al., 2003), verdicts (Bryant and Guilbault, 2002), experimental results (Slovic and Fischhoff, 1977), entrepreneurial success (Bukszar and Connolly, 1988), or how early a patient's condition could have been diagnosed (Arkes et al., 1981; Berlin, 2003).[1] A key mechanism of information projection is that such "creeping determinism" leads to systematically biased performance evaluations: evaluators who judge with hindsight systematically underestimate the competence of others.

We apply information projection to a competitive setting, in which the evaluators' ability matters—relative to those evaluated. Think about, for example, a prospective entrepreneur considering entering a market, co-workers considering competing for promotion, or a politician considering running for a new office. In all of these examples, the agents have no experience with the actual position they are looking to fill; yet they can judge some of their

[1] We restrict attention to hindsight bias only for the sake of exposition—many detrimental effects of information projection occur without reference to timing—i.e., in the form of the curse of knowledge. For example, knowing our intentions and what we want to say, we overestimate how clearly we come across when communicating with others (Newton, 1990; Keysar and Henly, 2002; Kruger et al., 2005). Hindsight bias is sometimes thought of as a within-person (across time) version of the curse of knowledge, albeit lots of the evidence on hindsight bias comes from between-subject settings (see, e.g., Fischhoff, 1975). We do not sharply distinguish between the two (both are fully covered by information projection) and use the terms synonymously.

competitors' performance from a hindsight perspective. For example, the prospective entrepreneur has to evaluate her skills relative to those of potential competitors. And she typically does so under informational asymmetries. She knows the demand for previously released products and services; she knows whether an investment opportunity turned out to be profitable; and she knows whether a hiring decision was "wise." With hindsight bias, she cannot disregard such ex post information when considering her competitors at the time of their decision making. In turn, she underestimates the uncertainty that they faced when making their decisions, and she overestimates the ex-ante probability of profitable managerial decision making and entrepreneurial success. With such inflated expectations, she is surprised when she sees that her competitors struggled and mistakenly ascribes this to their lack of talent. She believes that she can do better and, thus, competes too often (or enters prematurely). Such hindsight-biased overentry is highly inefficient because it affects low-skilled agents disproportionately more than it does high-skilled agents.

We test these predictions in an experiment in which subjects work on real-effort tasks and choose whether to compete against others or to be paid based on their individual performance (Niederle and Versterlund, 2007). In our setting, subjects see example tasks and learn about their competitors' past performance before choosing their payment scheme. In the *Hindsight treatment*, subjects know how to solve the example tasks. In the *No-Hindsight treatment*, subjects do not know the solutions to the example tasks. Otherwise, the two treatments are exactly the same.

The experiment confirms the main predictions. First, subjects in the Hindsight treatment, but not in the No-Hindsight treatment, overestimate the average success rate in the experimental task, i.e., how well anyone can do on the task. This generic overestimation is a direct manifestation of hindsight bias. Next, as predicted by information projection, subjects in the Hindsight treatment engage in biased performance evaluation: after seeing how their designated competitors performed in the past, subjects in the Hindsight treatment become significantly more optimistic about their tournament rank. This dynamic is not observed in the No-Hindsight treatment, suggesting that participants in the Hindsight treatment, on average, perceive their competitors as *below*-average, which is in line with information projection being the underlying mechanism. When making their entry choice, 49% of the Hindsight participants state that they will rank first in the tournament, while only 15% in the No-Hindsight treatment do so. Accordingly, Hindsight participants exhibit a significantly higher willingness to compete than subjects in the No-Hindsight treatment—the average tournament valuation in the Hindsight treatment is 19% higher than in the No-Hindsight treatment. The data also confirms that hindsight bias affects mainly low-performing subjects. The additional tournament entry in the Hindsight treatment is driven entirely by participants who should not have entered the tournament.

Earlier studies have identified predictors and mechanisms of comparative overconfidence and overentry. Many of these studies, like ours, add to the growing literature on the interplay between cognitive biases. Camerer and Lovallo (1999) show that overplacement and excessive market entry occur more often in environments with skill-based self-selection. With base-rate neglect, people ignore such self-selection and, therefore, underestimate the average talent of their competitors. Overconfidence can be the result of motivated reasoning, such as when "positive thinking" is performance-enhancing (Benabou and Tirole, 2002; Chen and Schildberg-Hörisch, 2019), when we derive "ego utility" from believing that we are competent (Kőszegi, 2006), or when overconfidence helps to persuade or deceive others (Schwardmann and van der Weele, 2019). Hindsight bias can also be motivated, for example, by a need for sensemaking, a desire to perceive the world in an orderly, predictable fashion, or, more closely related to ego utility, through direct pleasure from believing that we could have been able to foresee what happened (Roese and Vohs, 2012). Overconfidence can also arise through confirmatory bias (Rabin and Schrag, 1999): If we are too eager to interpret new information as supporting what we already believe, we will end up too certain about our views. Hindsight bias and confirmatory bias have in common that more information only exacerbates them; they can also reinforce each other in the sense that any initially overoptimistic belief is more likely to persist when confirmatory bias comes into play.[2] Finally, absent any actual overconfidence, a "culture of overconfidence" residing in higher-order beliefs can trigger behavior that is consonant with overconfidence (Bhaskar and Thomas, 2019).

In terms of structural predictors, Moore and Cain (2007) show that the degree of difficulty of a task predicts whether people are over- or underconfident. The easier a task, the more likely people are to be overconfident relative to others. Regarding individual characteristics, Niederle and Versterlund (2007) show that gender is a major predictor of overplacement. Men are more prone than women to overplacement and, therefore, enter competitive payment schemes more often (see, also, Dohmen and Falk, 2011). Several forms of overconfidence have been linked to personality traits, specifically, extraversion (Schaefer et al., 2004) and narcissism (John and Robins, 1994; Ames and Kammrath, 2004), suggesting that overconfidence may be, in part, a non-reducible bias. Our paper contributes to the research on the roots of overplacement and overentry by showing that these phenomena are more likely to occur in environments with informational asymmetries.

Another line of research related to our study investigates the economic consequences of information-projection biases. Camerer et al. (1989) show that the curse of knowledge distorts asset prices and that it survives in competitive market settings. Biais and Weber (2009) show that hindsight bias impedes learning about risks. Hindsight-biased investors are not surprised by asset price shocks and, therefore, do not sufficiently adjust their volatility estimates for future portfolio choices. Their field data shows that individual levels of hindsight bias among investment bankers are negatively related to their performance.

As in our study, Biais and Weber (2009) establish a connection between hindsight bias and a form of overconfidence. In their study, hindsight bias leads to overconfidence about the predictability of future outcomes, or *over-*

[2] We thank an anonymous referee for pointing this out.

*precision*, which is one of three main forms of overconfidence distinguished by Moore and Healy (2008). Here, people have too much faith in their estimates—their confidence intervals are too narrow. The second form of overconfidence is *overestimation*, which occurs when people overestimate their own performance levels. Hindsight bias itself can be seen as a specific form of overestimation. With outcome knowledge, we tend to overestimate how well we—and anybody else—could have predicted the state of the world. The third form of overconfidence is *overplacement*, which is the focus of this paper. Note that overestimation does not imply overplacement. First, from a theoretical perspective, a fully hindsight-biased decision maker thinks that anyone could have perfectly predicted the past and, thus, exhibits no overplacement at all. It is only through performance evaluation that hindsight bias can potentially lead to overplacement. Second, empirically, overplacement and overestimation are often *negatively* correlated. Moore and Healy's (2008) model of egocentric Bayesian learning explains this pattern through heterogeneity in task difficulty. Their model is relevant to our study, as part of our data suggests a hindsight-bias-driven "hard-*but-seemingly*-easy effect" that occurs without any information on performance.

Danz et al. (2015) relate hindsight bias to inefficient delegation. In their framework, hindsight bias manifests itself by principals having vague memories about their own past private signals, which leads them to overestimate their ability to decide correctly by themselves. Their framework also predicts comparative overconfidence. However, the authors do not observe or test beliefs in their delegation game. Apart from investigating entry into competitive tournaments, our study differs from theirs in two ways: (i) by providing a direct test of hindsight-biased performance evaluations in beliefs; and (ii) through our framework, which makes different predictions about the conditions under which overplacement occurs. In our framework, overplacement is the result of systematic underestimation of others' ability (through information projection) rather than of overestimation of one's own ability (through imperfect memory). This mechanical difference is important because it predicts overplacement even when the decision maker has no experience with the task (or no private information to misremember). As such, our paper is the first to provide a direct test of information projection as a mechanism for hindsight-biased performance evaluation.

In terms of metacognition, Loewenstein et al. (2006) show that people are not aware of their own hindsight bias and pay for information that biases their judgments and reduces their payoffs. People are, however, able to partially anticipate hindsight bias in others (Danz et al., 2018), which has been frequently suggested as an explanation for defensive agency practices such as overtreatment and the overuse of diagnostic tests by medical practitioners (Kessler and McClellan, 1996, 2002; Berlin, 2003; Studdert, 2005), as well as for managers' suboptimal levels of risk taking (Thaler, 2015). Our study adds to the literature on the economic consequences of hindsight bias by showing that it can cause overplacement and overentry into tournaments.

The paper is organized as follows. Section 2 provides a simple framework to pin down the predictions for our experiment, which is detailed in Section 3. Section 4 presents the results, and Section 5 concludes.

## 2 Theoretical framework

This section provides a simple application of Madarász's (2012) model of information projection to a market-entry problem. The purpose of this section is to clearly illustrate the effect of information projection on relative ability judgments and willingness to compete and to pin down our experimental hypotheses. We consider a simple inference problem of an agent who decides whether to compete against others after observing some of her competitors' past performance.

### Setup

Consider a market with two periods $t = 1, 2$. The incumbent $I$ serves the market in both periods. The entrant $E$ observes the incumbent's performance in the first period and can enter the market in the second period. We focus on $E$'s entry choice and abstract away from strategic uncertainty by assuming that the incumbent is known to have no incentive to leave the market in either period.

In each period, the state of the world $\omega_t$ is either $A$ or $B$. Nature draws the states independently across periods with known probability $P(A) = P(B) = 1/2$. Agents $i \in \{I, E\}$ in the market have to make a prediction $a_t^i \in \{A, B\}$ about the state based on a private signal $s_t^i \in \{A, B\}$. The agents differ in managerial talent $\theta_i$, which corresponds to the quality of their private signal $P(s_t^i = \omega_t | \theta_i) = \theta_i \geq 1/2$. That is, more-talented agents have a higher chance of reading the state of the world correctly. The entrant's talent $\theta_E$ is drawn from a continuous distribution with full support on $[0.5, 1]$ and cumulative distribution function $F(\theta)$. The incumbent's talent $\theta_I$ is either low or high, $\theta_I \in \{\underline{\theta}, \bar{\theta}\}, 0.5 \leq \underline{\theta} < \bar{\theta}$, with prior probability $\pi_0 := P(\theta_I = \bar{\theta}) \in (0, 1)$. The entrant knows her own talent but not the incumbent's talent; she knows only the prior $\pi_0$ of the incumbent's type. At the end of each period, the state of the world is revealed through signal $r_t \in \{A, B\}$, $P(r_t = \omega_t) = 1$.

In the first period, the incumbent is alone in the market. When making her entry choice in the second period, the entrant knows whether the incumbent successfully predicted the state of the world in period 1.[3] Let $m_0 > 0$ be the entrant's payoff in period 2 if she does not enter the market. If she enters the market, her payoff is $m_1 > m_0$ if she outperforms the incumbent—i.e., if she predicts $\omega_2$ correctly while the competitor does not—and zero otherwise.

In the following, we assume that the model parameters $\{\pi_0, \underline{\theta}, \bar{\theta}, m_0, m_1\}$ are chosen such that the market entry threshold (1) introduced below is in the interior of $(0.5, 1)$. This restriction excludes cases in which all entrants or no entrant would enter after observing a successful incumbent, which ensures that all predictions hold in the strict sense, and not just weakly. Figure 1 summarizes the entrant's inference problem.

[3] We assume that incentives are known to be that the incumbent always strictly prefers to predict the state of the world correctly.

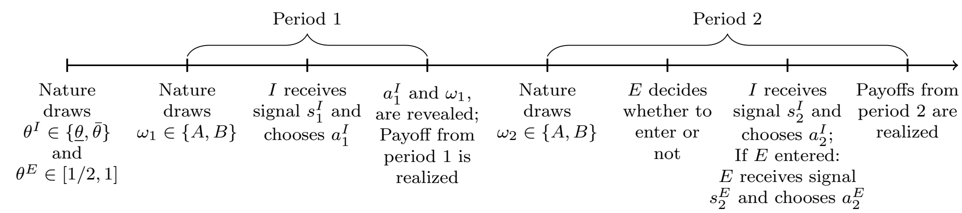


Figure 1: Timeline of the market-entry problem.

In the following, we compare a hindsight-biased entrant and an unbiased entrant to make clear how hindsight bias distorts relative performance evaluations and market entry.

**Unbiased entrant**

The inference of the unbiased entrant is as follows. First, her posterior about her opponent's type after observing her opponent's success $S\colon a_1^I = \omega_1$ is given by $\pi_S = \pi_0\bar{\theta}/(\pi_0\bar{\theta} + (1-\pi_0)\underline{\theta})$, and she enters if her talent exceeds a cutoff value that, under risk neutrality, is given by[4]

$$\theta_S^* = \frac{m_0}{m_1(\pi_S(1-\bar{\theta}) + (1-\pi_S)(1-\underline{\theta}))}. \tag{1}$$

The ex ante probability of market entry after observing a successful incumbent is, therefore, $1 - F(\theta_S^*)$. The threshold $\theta_{\overline{S}}^*$ for market entry after observing an unsuccessful incumbent $\overline{S}\colon a_1^I \neq \omega_1$ is derived accordingly, and the unconditional probability of market entry is

$$P(\text{Enter}) = P(S)(1 - F(\theta_S^*)) + (1 - P(S))(1 - F(\theta_{\overline{S}}^*)),$$

where $P(S) = \pi_0\bar{\theta} + (1-\pi_0)\underline{\theta}$ is the ex ante probability of observing a successful incumbent.

[4]Her chance of winning is $\theta_E(\pi_S(1-\bar{\theta}) + (1-\pi_S)(1-\underline{\theta}))$. She enters if her expected payoff from market entry is greater than her outside option, $E(m|S, \text{Enter}) = \theta_E(\pi_S(1-\bar{\theta}) + (1-\pi_S)(1-\underline{\theta}))m_1 \geq m_0$. As mentioned above, we focus on cases with $\theta_S^* \in (0.5, 1)$, in which some entrants choose to enter and some choose to not enter after observing a successful incumbent.

### Hindsight-biased entrant

Now consider an entrant with hindsight bias. When making her entry decision in period 2, the entrant has more information than the incumbent had in period 1—she knows the state of the world in period 1. The hindsight-biased entrant acts as if the incumbent somehow could have known the state when making her prediction.

We follow Madarász (2012) in modeling the entrant's hindsight bias.[5] Let $\Sigma_t^i$ be agent $i$'s information set at time $t$ and $p_t^i(\sigma_k) \in [0,1]$ be the true probability that she observes the realization of signal $\sigma_k$ in period $t$. At time $t'$, agent $j$ with information-projection bias of degree $\rho \in [0,1)$ perceives agent $i$'s information distribution at time $t$ as

$$\tilde{p}_t^i(\sigma_k) = \begin{cases} (1-\rho)p_t^i(\sigma_k) + \rho & \text{if } \sigma_k \in \Sigma_{t'}^j \\ p_t^i(\sigma_k) & \text{if } \sigma_k \notin \Sigma_{t'}^j \end{cases}. \tag{2}$$

That is, if $j$ has access to information $\sigma_k$ but others don't, then her perceived probability that others share(d) this information is a $\rho$-weighted average of the true probability and one. With $\rho = 0$, she is unbiased and perceives the distribution of information correctly. With $\rho \to 1$, she is fully biased and believes that all of her information is shared with everyone else.

In our setting, information projection can occur after the entrant has learned the state of the world in period 1. If the entrant projects the state-revealing signal $r_1$ onto the incumbent's information set in period 1, she will overestimate the ex ante success rate of any type of incumbent.[6] Specifically,

[5] Camerer et al. (1989) were the first to formalize the idea of hindsight bias and curse of knowledge. They capture the curse of knowledge by directly modeling expectations about outcomes as the $\rho$-weighted average of expectations with and without additional information. Biais and Weber (2009) use the same approach to show that hindsight bias impedes learning about risks. For a discussion of the differences between the models, see Madarász (2012).

[6] The incumbent's information set when making her prediction in period 1 is $\Sigma_1^I = \{s_1^I\}$, and the entrant's information set at the time of her entry choice in period 2 is $\Sigma_2^E = \{a_1^I, r_1\}$. Thus, the entrant's perceived probability that the incumbent had access to signal $r_1$ is $\tilde{p}_1^I(r_1) = (1-\rho)p_1^I(r_1)+\rho = \rho$ (the true probability $p_1^I(r_1)$ that the incumbent observed $r_1$ in $t = 1$ is zero). Her perceived chance of observing a successful incumbent

from the perspective of the biased entrant, the ex ante probability of seeing a successful incumbent in period 1 is

$$\widetilde{P}(S) = \pi_0[(1-\rho)\bar{\theta} + \rho] + (1-\pi_0)[(1-\rho)\underline{\theta} + \rho] > P(S), \rho > 0,$$

which is larger than that of the unbiased entrant and increasing in the degree of information projection $\rho$.[7] Note that there is no overplacement at this point. The entrant overestimates how well anyone could have performed—she does so to the same extent for herself as for others—and this mistake does not (yet) affect her inference about the incumbent's talent or her chance to win. Overplacement emerges when the entrant evaluates the incumbent's performance in hindsight. Specifically, when the hindsight-biased entrant observes that the incumbent predicted the state correctly, she is *less* surprised than the unbiased entrant. In turn, her posterior belief $\widetilde{\pi}_S$ about the talent of the incumbent after seeing her succeed is too conservative:[8]

$$\widetilde{\pi}_S = \frac{\pi_0[(1-\rho)\bar{\theta} + \rho]}{\widetilde{P}(S)} < \pi_S, \rho > 0.$$

In settings like ours, with relative performance evaluation, this underestimation of the others leads to comparative overconfidence. The entrant overestimates her chance of winning when facing a successful competitor and enters for a range of $\theta \in [\widetilde{\theta}^*_S, \theta^*_S)$,

$$\widetilde{\theta}^*_S = \frac{m_0}{m_1(\widetilde{\pi}_S(1-\bar{\theta}) + (1-\widetilde{\pi}_S)(1-\underline{\theta}))} < \theta^*_S,$$

where the unbiased entrant would not enter.[9]

given that the incumbent is of type $\theta^I$ is, therefore, $(1-\rho)\theta^I + \rho$ (knowing $r_1$ allows any type $\theta_I \in \{\underline{\theta}, \bar{\theta}\}$ to predict the state $\omega_1$ perfectly).

[7] Note that the biased entrant correctly perceives all prior distributions. The bias occurs because, due to projecting information, the entrant overestimates how well *any* type can do (Madarász, 2012).

[8] This is the case of "underinference" in Madarász (2012).

[9] Her biased estimate of her chance of winning is $\theta_E(\widetilde{\pi}_S(1-\bar{\theta}) + (1-\widetilde{\pi}_S)(1-\underline{\theta}))$. She enters if her perceived expected payoff from market entry is greater than her outside option, $\widetilde{E}(m|S, \text{Enter}) = \theta_E(\widetilde{\pi}_S(1-\bar{\theta}) + (1-\widetilde{\pi}_S)(1-\underline{\theta}))m_1 \geq m_0$, $\widetilde{E}(m|S, \text{Enter}) >$

The inference after observing a failing incumbent is not distorted in our simple setting since the ex post information $r_t$ always perfectly reveals the state $\omega_t$. Thus, on average—over failing and successful incumbents—the biased entrant underestimates the incumbent's talent.[10] This systematic underestimation is the key mechanism of Madarász's (2012) model of information projection tested in this paper. Since the entrant knows her own talent in our setting, her underestimation directly translates into comparative overconfidence.

**Prediction 1** (Overplacement)**.** *Hindsight-biased agents are more likely than unbiased agents to overestimate their ability relative to others'.*

Accordingly, the ex ante probability of market entry is higher with hindsight bias than in the unbiased case,[11]

$$\widetilde{P}(\text{Enter}) = P(S)(1 - F(\widetilde{\theta}^*_S)) + (1 - P(S))(1 - F(\theta^*_{\overline{S}})) > P(\text{Enter}).$$

**Prediction 2** (Increased market entry)**.** *Hindsight-biased agents are more likely than unbiased agents to enter competitive payment schemes.*

Finally, due to the shift in the entrant's cutoff, the expected talent of entrants to the market is decreasing in the level of information projection bias. In our simple setting, the additional entry comes entirely from participants who should not enter.

**Prediction 3** (Overentry)**.** *The average talent of market entrants is decreasing in the degree of hindsight bias.*

The above predictions are formulated as a comparison between the hindsight-biased and the unbiased agent. In our experiment, we cannot manip-

---

$E(m|S, \text{Enter}), \rho > 0$.

[10] With imperfect ex post information, the entrant's inference is also biased after observing a failing incumbent. Madarász (2012) shows that underestimation of others, on average (here, over failing and successful incumbents), holds as long as the outcome process (i) depends on the observee's competence and information; and (ii) satisfies the monotone likelihood ratio property in the observee's competence and her expected utility.

[11] The strict inequality holds because we focus on interior cases for (1) and because of the assumption that the distribution of the entrant's talent has full support on $[0.5, 1]$.

ulate the extent to which a person exhibits hindsight bias, and we make no assumptions about subjects' degree of hindsight bias, except that the distribution of hindsight bias is the same in both treatments (per randomization).

What we manipulate in the experiment is whether hindsight bias can affect a subject's inference. Our theoretical framework maps into the experiment as follows. The basic setup above corresponds to the treatment in which hindsight bias can affect a subject's inference. Our control treatment (no hindsight) corresponds to a modification of the above setup, in which the entrant does not learn the state of the world in period 1. Therefore, she cannot project any information onto the incumbent's information set. In this modified setting, even fully hindsight-biased agents make no mistakes and are behaviorally indistinguishable from the fully rational, unbiased agent. Thus, all of the above predictions can also be read as a comparison between a hindsight setting and a no-hindsight (control) setting.

# 3 Experimental Design

The experiment consisted of three treatments, with each subject participating in only one treatment. The experiment was programmed with z-Tree (Fischbacher, 2007). The experimental sessions were run at the Technische Universität Berlin in fall 2012. Participants were recruited using ORSEE (Greiner, 2004).

## 3.1 Tournament Treatment

We first conducted a *Tournament treatment*. We can think of the participants in this treatment as the "incumbents" in the toy model.

Figure 2: Example of a change-detection task. The images were presented alternately for 14 seconds on the computer screen (one second per image with 150ms interruptions).

In the Tournament treatment, subjects worked on 60 change-detection tasks (Rensink et al., 1997; Simons and Levin, 1997; Loewenstein et al., 2006). In each task, they had to spot the difference between two nearly identical images (see Figure 2).[12] Each image pair was presented in a short video clip of 14 seconds, with the two images displayed alternately with short interruptions. Then, subjects had 40 seconds to indicate the location of the difference. To this end, one of the images was displayed with a numbered grid, and subjects could enter one of the grid numbers (see Appendix C).

We chose this type of task because it was used as a hindsight stimulus beforehand (Loewenstein et al., 2006) and because it allows a precise quantification of hindsight bias. As a visual detection task, it resembles some aspects of diagnostics in radiology, where hindsight bias is known to have detrimental effects on legal decision making in malpractice claims (Berlin, 2003; Harley, 2007). However, hindsight bias has been demonstrated with many other tasks and contexts inside and outside of the lab, such as logical puzzles, trivia questions, prediction tasks, and case histories (see Christensen-Szalanski and Willham, 1991; Guilbault et al., 2004 for meta studies).

[12] Some of the tasks were generously provided by Rensink et al. (1997). We adopted the tasks in their "marginal interest" condition (except for those in which colors of objects changed) and designed the remaining tasks in the same fashion.

The 60 tasks were grouped into six rounds of ten tasks each.[13] At the beginning of each round, subjects were randomly matched to groups of four. Subjects earned 1 Euro for each task they solved if they ranked first relative to their matched participants and earned zero otherwise (ties were broken randomly). At the end of the experiment, two of the six rounds were randomly selected for payment. In addition to the payments from the tournaments, subjects received a 5-Euro show-up fee and a fixed payment of 7 Euro for working on the tasks. No feedback was provided at any time during the experiment.

Ninety-two subjects participated in the Tournament treatment in four sessions. The average duration of the sessions was 80 minutes and the average earning was 15.39 Euro.

### 3.2 Hindsight Treatment

The sessions for the two main treatments followed the completion of all the sessions of the Tournament treatment. The participants in these treatments mirror the "entrant" in the theoretical framework. Figure 3 gives an overview of the timeline of the *Hindsight treatment* and the *No-Hindsight treatment*.

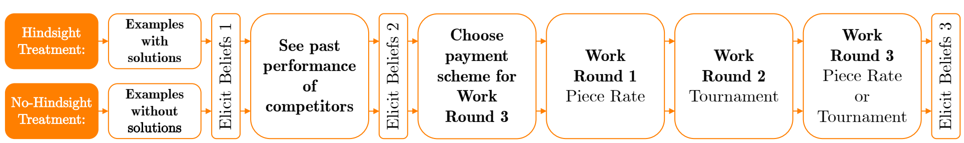


Figure 3: Timeline in the Hindsight treatment and the No-Hindsight treatment (treatments were varied between sessions).

At the outset of the two main treatments, participants learned about the procedures in the Tournament treatment and that they were being matched to three participants from the Tournament treatment.[14] Participants were

[13]Additionally, six example tasks were given at the outset of the experiment to familiarize subjects with the task.

[14]Matching subjects to past performances of other participants avoids potential con-

informed that this assignment would remain the same for the entire experiment.

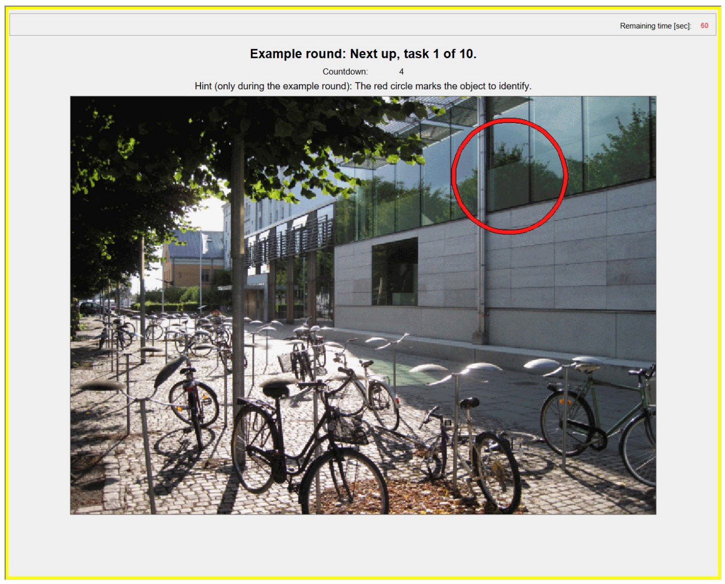


Figure 4: Solution to example task provided in the Hindsight treatment (translated from German).

Subjects then participated in an example round in which they saw ten different change-detection tasks taken from the Tournament treatment. In the *Hindsight treatment*, subjects received the solution to each example task: they saw the solution during a countdown announcing the next example task (see Figure 4).[15] The instructions (see Appendix C) and all ten solution screens in the example round (see Figure 4) made clear that they would receive the solutions only in the example round and not in the payoff-relevant

founds with other-regarding preferences or biased expectations about entry decisions of other players. To increase the power of our tests, we used a conditional random matching procedure in which subjects were matched based on specific performance patterns (see below). This procedure does not affect the inference a participant in the main treatment would rationally make if she saw the same performance pattern with unconditional random matching.

[15]This was done to make sure that subjects see the tasks in the example round in exactly the same way as their matched participants. See Appendix B for the original screenshots.

work rounds.[16] Apart from providing the solutions, the example round differed from the payoff-relevant rounds only in that subjects received no payment and that new tasks were taken from the Tournament treatment for each payoff-relevant round (subjects never saw the same task twice).

After the example round, subjects saw the performance of their matched participants on the ten example tasks. Specifically, they saw how many of their example tasks each of their matched participants solved (see Figure 5).[17] At this point, subjects were reminded that the tasks in their example round were not example tasks for their matched participants—i.e., that their matched participants performed these ten tasks in paid tournaments.

Next, subjects stated their willingness to compete with their matched participants in the third of the three payoff-relevant rounds (see Niederle and Versterlund, 2007). Each of the three payoff-relevant rounds consisted of ten new change-detection tasks taken from the Tournament treatment. In the first round, subjects were paid a piece rate. The piece rate was either 0.50

[16]It was important in our setting that participants did not mistakenly believe that the solutions were also available in the paid work rounds. Note that this could happen only if a subject did not understand the instructions and also missed all ten reminders during the example round, each explicitly pointing out that the solutions were available only during the example round (all belief elicitation and the entry choice took place after all rounds of the example round were concluded). If this was, nevertheless, the case, then we would expect to see a prominent mass of Hindsight participants predicting that they could solve ten out of ten tasks in the tournament. However, the distribution in the Hindsight treatment was unimodal, around 7/10 tasks (see right panel of Figure 8 in the Appendix) and there were only six participants in the Hindsight treatment and one participant in the No-Hindsight treatment who thought they could solve all tasks in the tournament. The additional participants in the Hindsight treatment appeared to be well in line with the shift in the belief distribution that was expected by information projection (see Figure 8 in the Appendix). Excluding all participants who thought they could solve all tasks in the tournament does not change our results qualitatively (see Tables 6 and 7 in the Appendix).

[17]The matching of participants in the main treatments to participants in the Tournament treatment was controlled and included only the most frequent performances. This was done to limit variation in observed performances and to ensure that our inference about optimal entry choices in the main treatments would be based on sufficient data. A third of the subjects observed that their three matched participants solved four, five, and five of the ten example tasks, respectively. Another third of the subjects observed that their matched participants solved four, five, and six of the ten example tasks, respectively. The remaining subjects observed that their matched participants solved five, five, and six of the ten example tasks, respectively.

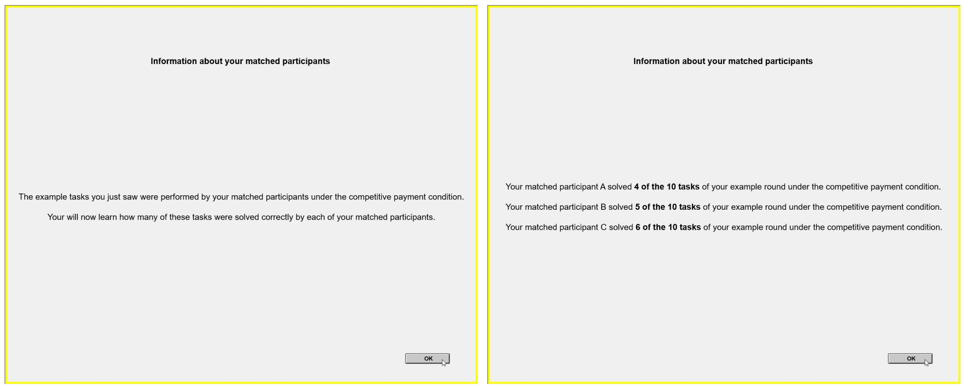


Figure 5: Information about matched participants (translated from German). The reminder (left panel) was shown before the performance data (right panel).

Euro or 0.54 Euro and was fixed within a session.[18] In the second round, subjects competed against their matched participants in a tournament. The tournament paid 2 Euro per correct task if the participant ranked first, i.e., if she solved more tasks than each of her matched participants (ties were broken randomly), and 0 Euro otherwise. In the third round, subjects were paid either according to the piece rate or according to the tournament, depending on their elicited willingness to compete.

We used a variation of the Becker-DeGroot-Marschak mechanism (see Bohnet and Zeckhauser, 2004) to elicit subjects' willingness to compete. Subjects stated their *piece-rate equivalent* knowing that the actual piece rate would be announced after all subjects stated their piece-rate equivalent. A subject was paid according to the actual piece rate if it turned out to be at least as high as her piece-rate equivalent. Otherwise, she was paid according to the tournament.[19,20]

[18] We chose the piece rates (i) such that a perfectly calibrated subject with an average performance level would be indifferent between the piece rate and the tournament; and (ii) to allow a comparison of our data with that of Niederle and Versterlund (2007).

[19] This mechanism is incentive-compatible with truth-telling being strictly dominant as long as subjects assign positive subjective probabilities to piece rates in the neighborhood of their piece-rate equivalent (see Bohnet et al., 2008).

[20] To prevent subjects from speculating that their choice of the minimum piece rate

We elicited subjects' beliefs three times: after the example round; after they saw the performance of their matched participants; and at the end of the experiment. All belief elicitation in the experiment referred to the tournament in round 2. We use the first and the second belief elicitation to test predictions for the experiment.[21] At each elicitation, we asked subjects to make three guesses. First, we asked them to guess the number of tasks that a randomly selected other participant of the same session would solve in the tournament. This measured subjects' perception of the average success rate on the task—i.e., their perceived degree of difficulty of the task. Second, we asked them to guess the number of tasks *they* would solve in the tournament. Finally, we asked them to guess their rank in the tournament. We used this guess to test for hindsight-biased overplacement. At the end of the experiment, one of the stated beliefs was randomly selected for payment. Subjects earned 1 Euro if their answer was correct and 0 otherwise (see Niederle and Versterlund, 2007).

We elicited subjects' risk attitudes at the beginning of the sessions, using a multiple price list similar to that of Holt and Laury (2002). In 21 cases, the subjects had to decide whether they preferred a safe payment or a lottery. In each case, the lottery paid 6 Euro or 0 Euro with equal probability. The safe payoff was increased in each case from 0 to 6 Euro in increments of 0.30 Euro. One of their choices was randomly selected for payment at the end of the experiment.

At the end of the experiment, subjects completed a questionnaire on basic sociodemographics, a brief version of the Big-Five Personality Test (Rammstedt and John, 2007), the Life Orientation Test (LOT-R; Glaesmer et al., 2008), and a seven-item questionnaire on risk attitudes from the German Socio-Economic Panel (SOEP).

Subjects received feedback about their own performance only at the end

might affect the piece rate, the piece rate was stowed in a sealed envelope and attached to the front door of the lab (see Bohnet and Zeckhauser, 2004).

[21] The third belief elicitation differs from the other two elicitations by being retrospective rather than prospective. We discuss the results of the third belief elicitation in the final section of the paper.

of the experiment, when they learned their absolute performance and their rank in each round, the randomly selected payment round, and their payoffs for the change-detection task and the risk and belief elicitation tasks. Then, subjects were guided to the next room, where they received their payment in private.[22]

### 3.3 No-Hindsight Treatment

The *No-Hindsight treatment* was exactly the same as the Hindsight treatment—except for the example round. Participants in the No-Hindsight treatment did not receive the solutions to the example tasks (the countdown screen in Figure 4 did not show the image with the red circle). Thus, subjects had no extra information to project that could potentially distort their inference and entry choice.

We ran four sessions for the Hindsight treatment and four sessions for the No-Hindsight treatment, each with 18 subjects. Invitations to the sessions were gender-balanced. The average duration of the sessions was one hour and 44 minutes, and the average payoff was 21.13 Euro.

### 3.4 Experimental Hypotheses

The experimental hypotheses closely follow the predictions above. Like the entrant who evaluates competitors with ex post information, the participants in the Hindsight treatment know how to solve the tasks when assessing the ability of their matched participants (henceforth, *competitors*). They will overestimate the average success rate in the change-detection tasks if they mistakenly project this information, which, in turn, can lead to overconfidence relative to their competitors when they see how their competitors performed.

[22]To avoid curiosity effects, the instructions made clear that subjects would receive feedback about their performance level and rank in each round, regardless of whether or not they ended up in the tournament in round 3.

**Hypothesis 1** (Overplacement)**.** *After seeing their competitors' performance, participants in the Hindsight treatment, but not those in the No-Hindsight treatment, overestimate their rank in the tournament.*

The overly optimistic expectations about tournament ranks directly translate into more tournament entry.

**Hypothesis 2** (Increased tournament entry)**.** *Participants in the Hindsight treatment enter the tournament more often than participants in the No-Hindsight treatment.*

Finally, we hypothesize that the additional tournament entry in the Hindsight treatment comes mostly from participants with suboptimal ability levels. This follows not only from our toy model, but also from a simple ceiling-effect logic if participants also learn about their *own* ability when going through the example tasks. After seeing the example tasks, a top performer will correctly expect to solve most tasks in the tournament. Information projection cannot bias her expectations much further since her true success rate is close to one. In contrast, a participant with a success rate close to zero offers plenty of room for hindsight bias to create the illusion of doing well on the task.

**Hypothesis 3** (Overentry)**.** *The average ability of tournament entrants is lower in the Hindsight treatment than in the No-Hindsight treatment.*

# 4 Results

The left panel of Figure 6 shows the effect of hindsight bias on subjects' perceived degree of the tasks' difficulty after seeing the examples (first belief elicitation). As expected, participants in the Hindsight treatment had significantly higher expectations about the average success rate (66.3%) than did subjects in the No-Hindsight treatment (50.3%; $p < 0.001$, rank-sum test). In the Hindsight treatment, subjects significantly overestimated the true success rate of others on the tasks ($p < 0.001$, one-sample $t$-tests), while participants in the No-Hindsight treatment were, on average, well calibrated ($p = 0.869$).[23] The generic overoptimism in the Hindsight treatment is a direct manifestation of the hindsight bias. Remember that at this point—when subjects had just gone through the examples—hindsight bias was not expected to generate any overplacement. Overplacement was predicted only when hindsight-biased participants saw their competitors' performance.

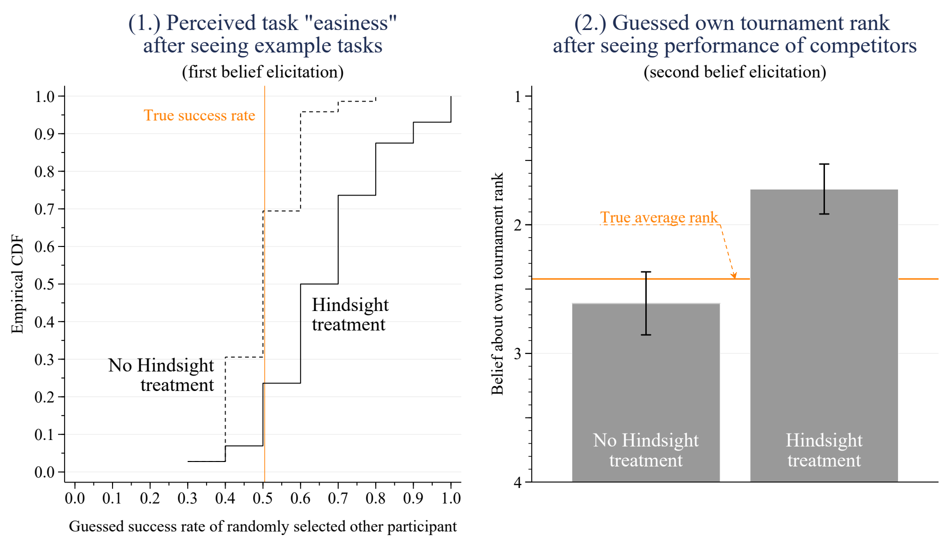


Figure 6: Guessed average success rate after seeing example tasks (left panel) and average estimate of own tournament rank after seeing performance of competitors (right panel). Bars represent 95% confidence intervals.

The right panel of Figure 6 shows the average anticipated tournament rank in each treatment after subjects saw the past performance of their competitors (second belief elicitation). In line with Hypothesis 1, after observing their competitors' performance, participants in the Hindsight treatment were significantly more optimistic about their tournament rank than were participants in the No-Hindsight treatment ($p < 0.001$, rank-sum test). As predicted, this overplacement was driven (partially) by performance evalu-

[23] We apply one-sample $t$-tests whenever the location parameter of a distribution is tested against a constant. In all cases, the results of one-sample signed-rank tests are qualitatively the same.

ation: When comparing subjects' beliefs before and after seeing the performance of their competitors (first versus second belief elicitation), we found that subjects in the Hindsight treatment became more optimistic about winning the tournament ($p = 0.045$, signed-rank test). This finding suggests that they were surprised by their competitors' low performance levels and that they perceived their competitors as below-average performers. In fact, in the Hindsight treatment, participants' estimate of the overall success rate on the task remained significantly higher than what they saw for the past performance of their competitors ($p < 0.001$, signed-rank test). However, this difference was much smaller after they saw the performance data.[24] The same dynamics were not observed in the No-Hindsight treatment. In fact, here, the pattern went partially in the opposite direction.[25]

The magnitude of the treatment effect on overplacement after performance evaluation is large and also significant if we control for performance by testing individual differences between guessed ranks and actual tournament ranks between treatments ($p = 0.001$, rank-sum test).[26] While par-

[24]In a previous version of this paper (Danz, 2014), we find that participants in the Hindsight treatment show no significant hindsight bias after observing the past performance of their competitors. Our test here is conceptually slightly different and has more power since we are testing individual performance estimates against individual averages over the past performances that a participant observes (rather than against the overall actual success rate on the task).

[25]The No-Hindsight treatment participants exhibit more *under*placement after seeing their competitors' past performance ($p = 0.032$). Their estimates of the average success rate on the task are not different from the average past performance of their competitors' they observe, both before and after they see their competitors' past performance ($p = 0.872$ and $p = 0.110$ from signed-rank tests, respectively). The change in guessed tournament ranks over time is significantly different between treatments ($p = 0.003$, rank-sum test).

[26]It is important in our setting that the solutions in the example round have no performance-enhancing effect in the work rounds. There are no significant performance differences between the treatments at the 5% level, on the aggregate or on the work round level. However, we find some indication of this effect. The average absolute score is somewhat lower in the No-Hindsight treatment than in the Hindsight treatment in the first round (4.4 versus 5, $p = 0.062$) and very similar in the tournament ($p = 0.492$) and in round 3 ($p = 0.654$). Subjects' average rank over all rounds is 2.5 in the No-Hindsight treatment and 2.4 in the Hindsight treatment (round-specific rank-sum tests between treatments yield $p = 0.271$, $p = 0.606$, and $p = 0.786$ for round 1, 2, and 3, respectively). In our analysis, we control for performance differences by running separate regressions for each work round with dummies for each performance level. The results are qualitatively the same, and effect sizes are not considerably different across these regres-

ticipants in the No-Hindsight treatment are well calibrated, on average, the guessed rank of participants in the Hindsight treatment is 1.7, on average, which is significantly higher than their true average rank or the expected rank of 2.5 ($p < 0.001$ each).[27]

**Result 1** (Overplacement)**.** *Participants in the Hindsight treatment, but not in the No-Hindsight treatment, (i) become more overconfident about winning the tournament after seeing their competitors' past performance; and (ii) overestimate their rank in the tournament, on average.*

We note that there is a smaller, but highly significant, treatment difference in guessed tournament ranks also *before* participants saw the performance of their competitors, i.e., right after they saw the example tasks (first belief elicitation, $p < 0.001$). This is not predicted by the basic framework but can be explained by a combination of hindsight bias and the Bayesian hard-easy effect (Moore and Healy, 2008). Suppose that a subject learns not only about her opponent's ability, but also about (i) her own ability and (ii) the degree of difficulty of the task (the average success rate, over all subjects). The subject has some prior belief about the average success rate and her own success rate, say 50% for each. Suppose that the task is actually easier and has a true success rate of 80%. Going through the examples, the subject imagines how many tasks she could have completed in a real, competitive situation. In the Moore and Healy (2008) model,

sions (see Tables 3 and 4 in the Appendix). The rank sum test above yields $p = 0.007$ and $p = 0.003$ when using ranks based on round 1 and round 3, respectively, instead of the tournament round. Danz et al. (2018) use the same hindsight stimuli as in our study but avoid this issue by showing the example task solutions also in the No-Hindsight treatment, *after* each example, such that subjects can see each task from the ex ante perspective (like the incumbent), but also benefit from any potential informational value of seeing *additional* solutions (in the No-Hindsight treatment, subjects found 41% of the example task solutions). Their hindsight effect (17.7 percentage points) is even slightly larger than ours (16.0pp), which suggests that our treatment differences in beliefs reflect hindsight bias and not false anticipations of a performance-enhancing effect of seeing all solutions.

[27]The guessed rank of No-Hindsight participants is not significantly different from their true average rank ($p = 0.472$, signed-rank tests) or from the expected rank of 2.5 ($p = 0.369$, one-sample $t$-test).

such introspective inference is more informative about her own skill than about others'. That is, after seeing the unexpectedly easy examples, she would update her belief about her own success rate upwards, and—to a lesser extent—her belief about the average success rate. This asymmetry in learning results in overplacement when facing unexpectedly easy tasks. The reverse effect, underplacement, occurs in the presence of unexpectedly hard tasks.

Note that the key to any overplacement in our setting is hindsight bias, which alters how the participant *perceives* the task—i.e., whether the examples give her the impression of a comparatively hard or easy task. Hindsight bias distorts this impression by making tasks look easier than they actually are and thereby causes overplacement not only when observing others (as in the basic framework), but also through egocentric Bayesian learning based on the misperceived easiness of the task (a "hard-*but-seemingly*-easy effect" in the Moore and Healy, 2008 model). To summarize, hindsight bias is necessary to explain treatment differences in overplacement and overentry in our setting; the Bayesian hard-easy effect *alone* predicts the same level of over- or underplacement in both treatments, at any time, and, thus, cannot explain our findings.

We move on to test whether the treatment differences in relative performance evaluations translate to differences in tournament entry behavior. Figure 7 plots the empirical cumulative distribution functions of subjects' piece-rate equivalents for tournament participation in each treatment. The figure shows that the distribution in the Hindsight treatment first-order stochastically dominates the distribution in the No-Hindsight treatment. The average piece-rate equivalent is 0.95 Euro in the No-Hindsight treatment and 1.13 Euro in the Hindsight treatment ($p = 0.027$, rank-sum test). That is, hindsight increases participants' valuation of tournament participation by 19%.[28]

[28]Our experiment shows only that hindsight causes hindsight bias and that hindsight causes overplacement and overentry. The further interpretation of our results relies on the assumption that hindsight bias is the mediator between hindsight on the one hand and overplacement and overentry on the other hand, as laid out in the theoretical framework.

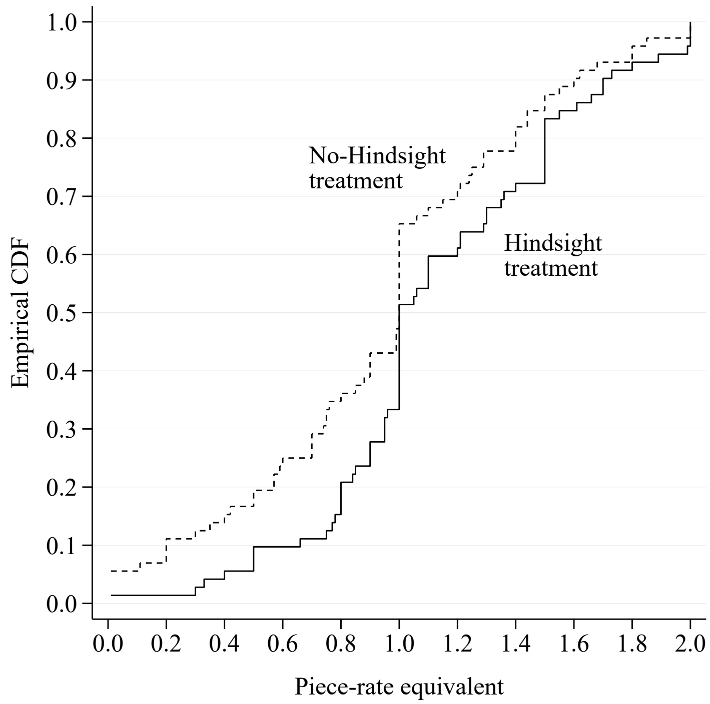


Figure 7: Distribution of individual piece-rate equivalents for tournament participation in the No-Hindsight treatment (dashed line) and the Hindsight treatment (solid line).

Table 1 shows regressions of the piece-rate equivalents on a treatment dummy, performance controls, and subjects' beliefs. The first two columns show that the size of the estimated treatment effect on the piece rate is very similar and significant when we control for individual performance levels. The third column shows that the treatment effect can be fully explained by subjects' relative performance expectations. Once we add subjects' guessed tournament ranks as a control, the treatment effect in piece-rate equivalents is insignificant and close to zero.[29] This finding shows that beliefs and

We are not aware of any mechanism that could explain our treatment differences by a direct causal link (without hindsight bias) between hindsight and overplacement and overentry, or through an alternative mediator that is unrelated to hindsight bias.

[29]There are no significant gender differences in the treatment effects on subjects' guessed tournament ranks or their piece-rate equivalents. We ran an ordered probit [tobit] regres-

actions were consistent in our setting and that the additional tournament entry in the Hindsight treatment was driven entirely by subjects' inflated expectations about their relative ability—as predicted by Hypotheses 1 and 2.

Table 1: Regressions of piece-rate equivalent on treatment, performance, and guessed tournament rank.

| Dependent variable: (Tobit regressions) | Piece-rate equivalent (1) | (2) | (3) |
|---|---|---|---|
| Hindsight treatment | $0.184^{**}$ | $0.180^{**}$ | 0.016 |
| | (0.079) | (0.077) | (0.080) |
| Guessed rank | | | $-0.181^{***}$ |
| | | | (0.039) |
| Dummies for performance levels | No | Yes | Yes |
| Constant | $0.958^{***}$ | $1.320^{***}$ | $2.026^{***}$ |
| | (0.056) | (0.460) | (0.456) |
| $N$ | 144 | 144 | 144 |
| $\log L$ | $-100.338$ | $-94.267$ | $-84.350$ |

Note: Columns 2 and 3 include dummies for each performance level in the tournament (round 2). The results are qualitatively the same when adding dummies for each performance level in round 1 or round 3 instead of round 2. (Guessed) tournament ranks range from 1 (best) to 4. Values in parentheses are standard errors: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

**Result 2.** ***(Increased willingness to compete)*** *(1) Piece-rate equivalents are significantly higher in the Hindsight treatment than in the No-Hindsight treatment. (2) The treatment difference can be fully explained by differences in subjects' expectations about their tournament ranks.*

Was the increased tournament entry driven primarily by participants with comparatively low ability, as suggested by Hypothesis 3? To investigate this question, we divided subjects according to their tournament performance $x$ and classified subjects with $x \geq 6$ as *high performers* and those

sion of subjects' guessed tournament ranks [piece-rate equivalents] on a treatment dummy, a gender dummy, and the interaction of both. Testing the interaction effect against zero yields $p = 0.660$ [$p = 0.944$].

with $x < 6$ as *low performers*.[30] We chose this specific threshold because it (i) yields a relatively balanced number of subjects in both groups and (ii) is informative about the optimality of subjects' tournament entry: for a given performance, an expected payoff maximizer is indifferent between the piece rate and the tournament if $w_P x = w_T P(R = 1 \mid y_0, x)x$, where $w_P$ denotes the piece rate; $w_T = 2$ denotes the payment per correct answer in the tournament in the event of winning; and $P(R = 1 \mid y_0, x)$ denotes the probability of winning the tournament conditional on the observed past performance of the competitors $y_0$ and one's own performance $x$. Given the wage parameters and the empirical winning probabilities, expected payoffs from tournament entry exceeded those with piece-rate compensation whenever a subject solved at least six tasks. Accordingly, we consider sorting choices as optimal if low performers sorted into the piece rate and high performers sorted into the tournament. We will refer to the reverse cases as *overentry* and *underentry*, respectively.

Table 2 shows regressions of tournament entry on the treatment dummy, run separately for high and low performers, as defined above. The table shows that the treatment effect is significant for low performers ($p = 0.029$) but not for high performers ($p = 0.775$). The treatment effect is significantly larger for low performers than for high performers if we control for individual risk attitudes, gender, and further variables (a one-sided test of the interaction effect yields $p = 0.032$; see Table 5 in the appendix).[31] These findings support Hypothesis 3: hindsight generates overentry.

**Result 3.** ***(Overentry)*** *The increased tournament entry in the Hindsight*

[30]The classification based on individual performance in the tournament (round 2) is significantly correlated with classifications based on round-1 and round-3 performances (chi-squared tests of independence yield $p = 0.029$ and $p = 0.001$, respectively). The share of participants that is classified the same way is 59.7% and 64.6%, respectively.

[31]We allow a one-sided test here since we test a directed ex ante hypothesis. The interaction effect is significant if we classify subjects as high and low performers according to their tournament (round 2) performance but not if the classification is based on their round-1 (piece rate) or round-3 (endogenous) performance. However, separate tests of the treatment effect for low- and high performers yield a consistent pattern throughout all classifications: the treatment effect is always significant for low performers, while it is never significant for high performers (all one-sided tests).

Table 2: Treatment effect on tournament participation conditional on performance level.

| Dependent variable: (Probit regressions) | Tournament entry | | | |
|---|---|---|---|---|
| | Low performers | | High performers | |
| | (1) | (2) | (3) | (4) |
| Hindsight treatment | 1.055** | 1.270** | 0.103 | −0.029 |
| | (0.484) | (0.602) | (0.359) | (0.415) |
| Risk taking (SOEP) | | 0.189* | | 0.135 |
| | | (0.107) | | (0.100) |
| Gender (1-male) | | 0.254 | | 0.671 |
| | | (0.441) | | (0.416) |
| High School grade (1-4,1-best) | | −0.673 | | 0.701* |
| | | (0.460) | | (0.416) |
| Age | | 0.055 | | 0.005 |
| | | (0.061) | | (0.044) |
| Constant | 0.859*** | 0.090 | 0.865*** | −1.584 |
| | (0.224) | (1.626) | (0.259) | (1.149) |
| $N$ | 77 | 77 | 67 | 67 |
| $\log L$ | −24.806 | −21.624 | −31.451 | −25.720 |

Note: Subjects were classified as low performers [high performers] if they solved fewer than [at least] six tasks in the tournament (round 2). Values in parentheses are standard errors: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

*treatment is driven by low performers: low performers enter significantly more often in the Hindsight treatment than in the No-Hindsight treatment; high performers enter at the same rate in both treatments.*

# 5 Discussion

This study presents experimental evidence that hindsight bias—a robust distortion of how humans process informational asymmetries—causes overconfidence in relative performance evaluations and overconfident behavior in the form of overentry into competition. In our setting, hindsight bias increases individual valuation of tournament entry by 19%, on average. As predicted, hindsight-biased tournament entry is driven by individuals who would have been better off not entering the tournament.

Our data shows an additional, unexpected hindsight effect that is independent of performance evaluation. Hindsight bias triggered comparative overconfidence in beliefs even before decision makers had seen any performance data (first belief elicitation). This immediate hindsight effect was not predicted by our basic framework but can be captured by a "hard-*but-seemingly*-easy effect"—a combination of hindsight bias and the Bayesian hard-easy effect (Moore and Healy, 2008). In hindsight, things are *seemingly* predictable and easy (although they were not), and seemingly easy tasks produce overplacement through egocentric Bayesian learning, in the same way as tasks that are actually relatively easy (see Moore and Healy, 2008). Together, our findings suggest that hindsight bias might be a powerful driver of overplacement in real-world settings—in the realm of performance evaluation, but also in settings without any information on performance.

Our experiment confirms information projection (Madarász, 2012) as a mechanism underlying hindsight bias and biased performance evaluations. There are various other testable implications of information projection and hindsight-biased performance evaluation, such as suboptimal information acquisition (Roese and Vohs, 2012) and scapegoating (Madarász, 2012). For example, the goal of mortality and morbidity conferences among clinical practitioners is to identify factors that led to complications or death and to review potentially unsafe clinical practices. With hindsight bias, this process might end prematurely since, more often than is warranted, a single explanation (such as incompetence) appears to fully account for the outcome, while other important factors (such as unclear responsibilities and communication channels) remain unnoticed and persist. At the same time, unclear or too little communication can be a direct result of information projection (Madarász, 2012). Information projection can also lead to favoritism and discrimination under heterogeneity in linguistic backgrounds or when people are subject to different levels of monitoring (Madarász, 2012). Finally, the anticipation of hindsight-biased performance evaluations can lead to "defensive agency," in which agents change the production of information to prevent reputational harm and false liability (Kessler and McClellan, 1996;

Madarász, 2012).

While we are not aware of any mechanism other than hindsight bias (or curse of knowledge) that can explain our findings, a limitation of our study is that it does not offer an explanation for persistent overconfidence in tournaments (see Huffman et al., 2019). In our setting, agents who learn about their own talent while gaining experience with the task become less overconfident over time. However, important instances of overconfidence, such as small-business failure, are marked by excess entry, followed by learning and exit, and might, therefore, be well explained by hindsight-biased overplacement.

What can be done if future research corroborates the importance of hindsight bias for overplacement and overentry? A few known moderators of hindsight bias might also help to reduce overplacement and overentry. First, individual expertise and familiarity with a task have been linked to lower levels of hindsight bias (Dawson et al., 1988; Christensen-Szalanski and Willham, 1991). While we cannot test this relationship in our setting, our data shows that hands-on experience with the task can correct hindsight-biased performance evaluation and may, thus, potentially reduce overentry. In retrospect—i.e., after performing the task but before receiving any feedback (third belief elicitation)—participants in the Hindsight treatment correctly assessed their performance relative to their competitors', on average, which suggests that they no longer believed that their competitors were "below average."[32] Acquiring management experience and business-related expertise before startup formation may, therefore, help to correctly assess the competitiveness of a startup and the odds of business success (see, also, Azoulay et al., 2020, who find that industry-specific experience

[32]In retrospect, participants also correctly assessed their own absolute performance and the absolute performance of a randomly selected other participant (whom, in contrast to their competitors, they had not monitored initially). However, these observations do not rule out that participants in the Hindsight treatment remained overoptimistic regarding their future performance or the easiness of the task since, for example, subjects might have believed that they got an "unlucky" sample of tasks in the tournament (all belief elicitation, including the third, referred to the tournament in round 2).

predicts entrepreneurial success).[33] Another known moderator of hindsight bias is forced deliberation: When people are asked to provide reasoning for their judgments in hindsight or potential explanations for counterfactual outcomes, hindsight bias tends to be weaker (Slovic and Fischhoff, 1977; Arkes et al., 1988; Kennedy, 1995). If such reasoning can be self-invoked in an analytical culture that is mindful of hindsight bias, biased performance evaluations and overentry might be mitigated.

On a final note, our findings may be of practical interest to experimenters studying overconfidence. Our setting provides a simple tool for exogenously manipulating comparative (over)confidence and the inclination to compete.

## References

Ames, Daniel R., Lara K. Kammrath. 2004. Mind-Reading and Metacognition: Narcissism, not Actual Competence, Predicts Self-Estimated Ability. *Journal of Nonverbal Behavior* **28**(3) 187–209.

Anderson, John C, Marianne M Jennings, David Lowe, Philip Reckers. 1997. The mitigation of hindsight bias in judges' evaluation of auditor decisions. *Auditing* **16**(2) 35–39.

Arkes, Hal R. 2013. The Consequences of the Hindsight Bias in Medical Decision Making. *Current Directions in Psychological Science* **22**(5) 356–360.

Arkes, Hal R., David Faust, Thomas J. Guilmette, Kathleen Hart. 1988. Eliminating the hindsight bias. *Journal of Applied Psychology* **73**(2) 305–307.

Arkes, Hal R, Robert L Wortmann, Paul D Saville, Allan R Harkness. 1981.

[33] Azoulay et al.'s (2020) data challenge the view that youth is a key factor in successful entrepreneurship. Among the 2.7 million founders in the US between 2007–2014 that started a business and had at least one hire, the average age at founding of the fastest-growing ventures was 45.

Hindsight bias among physicians weighing the likelihood of diagnoses. *Journal of Applied Psychology* **66**(2) 252–254.

Azoulay, Pierre, Benjamin F Jones, J Daniel Kim, Javier Miranda. 2020. Age and High-Growth Entrepreneurship. *American Economic Review: Insights* **2**(1) 65–82.

Barber, Brad M., Terrance Odean. 2001. Boys will be Boys: Gender, Overconfidence, and Common Stock Investment. *The Quarterly Journal of Economics* **116**(1) 261–292.

Benabou, R., J. Tirole. 2002. Self-Confidence and Personal Motivation. *The Quarterly Journal of Economics* **117**(3) 871–915.

Berlin, Leonard. 2003. Testimony to the {U.S.} Senate Committee on Health, Education Labor and Pensions: Mammography Quality Standards Act Reauthorization.

Bhaskar, V., Caroline Thomas. 2019. The Culture of Overconfidence. *American Economic Review: Insights* **1**(1) 95–110.

Biais, Bruno, Martin Weber. 2009. Hindsight Bias, Risk Perception, and Investment Performance. *Management Science* **55**(6) 1018–1029.

Blank, Hartmut, Volkhard Fischer, Edgar Erdfelder. 2003. Hindsight bias in political elections. *Memory* **11**(4-5) 491–504.

Bohnet, I, R Zeckhauser. 2004. Trust, Risk and Betrayal. *Journal of Economic Behavior & Organization* **55**(4) 467–484.

Bohnet, Iris, Fiona Greig, Benedikt Herrmann, Richard Zeckhauser. 2008. Betrayal Aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States. *American Economic Review* **98**(1) 294–310.

Bryant, Fred B., Rebecca L. Guilbault. 2002. "I Knew It All Along" Eventually: The Development of Hindsight Bias in Reaction to the Clinton Impeachment Verdict. *Basic and Applied Social Psychology* **24**(1) 27–41.

Bukszar, E., T. Connolly. 1988. Hindsight bias and strategic choice: Some problems in learning from experience. *Academy of Management Journal* **31**(3) 628–641.

Camerer, Colin, George F Loewenstein, Martin Weber. 1989. The Curse of Knowledge in Economic Settings: An Experimental Analysis. *Journal of Political Economy* **97**(5) 1232–1254.

Camerer, Colin, Dan Lovallo. 1999. Overconfidence and Excess Entry: An Experimental Approach. *American Economic Review* **89**(1) 306–318.

Cassar, Gavin, Justin Craig. 2009. An investigation of hindsight bias in nascent venture activity. *Journal of Business Venturing* **24**(2) 149–164.

Chen, Si, Hannah Schildberg-Hörisch. 2019. Looking at the bright side: The motivational value of confidence. *European Economic Review* **120** 103302.

Christensen-Szalanski, Jay J.J., Cynthia Fobian Willham. 1991. The hindsight bias: A meta-analysis. *Organizational Behavior and Human Decision Processes* **48**(1) 147–168.

Cooper, Arnold C., Carolyn Y. Woo, William C. Dunkelberg. 1988. Entrepreneurs' perceived chances for success. *Journal of Business Venturing* **3**(2) 97–108.

Danz, D, D Kuebler, L Mechtenberg, J Schmid. 2015. On the failure of hindsight-biased principals to delegate optimally. *Management Science* **61**(8) 1938–1958.

Danz, David N. 2013. The curse of knowledge increases self-selection into competition. *Three essays on irrational beliefs*. Dissertation. Technische Universität Berlin, Germany, 61–128.

Danz, David N. 2014. The curse of knowledge increases self-selection into competition: Experimental evidence. *WZB Discussion Paper, No. SP II 2014-207* .

Danz, David N, Kristóf Madarász, Stephanie W Wang. 2018. The Biases of Others: Projection Equilibrium in an Agency Setting. *Discussion Paper DP12867, CEPR.* .

Dawson, Neal V., Hal R. Arkes, Carl Siciliano, Richard Blinkhorn, Mark Lakshmanan, Mary Petrelli. 1988. Hindsight Bias: An Impediment to Accurate Probability Estimation in Clinicopathologic Conferences. *Medical Decision Making* **8**(4) 259–264.

Dohmen, Thomas, Armin Falk. 2011. Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender. *American Economic Review* **101**(2) 556–590.

Fischbacher, Urs. 2007. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* **10**(2) 171–178.

Fischhoff, Baruch. 1975. Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Judgment and Decision Making* **1**(3) 288–299.

Fischhoff, Baruch, Ruth Beyth. 1975. I knew it would happen. *Organizational Behavior and Human Performance* **13**(1) 1–16.

Gilovich, Thomas, Victoria Husted Medvec, Kenneth Savitsky. 1998. The Illusion of Transparency: Biased Assessments of Others' Ability to Read One's Emotional States. *Journal of Personality and Social Psychology* **75**(2) 332–346.

Glaesmer, Heide, Jürgen Hoyer, Jens Klotsche, Philipp Yorck Herzberg. 2008. Die deutsche Version des Life-Orientation-Tests (LOT-R) zum dispositionellen Optimismus und Pessimismus. *Zeitschrift für Gesundheitspsychologie* **16**(1) 26–31.

Glaser, Markus, Thomas Langer, Martin Weber. 2005. Overconfidence of Professionals and Lay Men: Individual Differences Within and Between Tasks? *Discussion Paper 05-25, SFB 504, University of Mannheim* 1–32.

Greiner, Ben. 2004. An Online Recruitment System For Economic Experiments. *Forschung und Wissenschaftliches Rechnen* **63** 79–93.

Guilbault, Rebecca L, Fred B Bryant, Jennifer Howard Brockway, Emil J Posavac. 2004. A Meta-Analysis of Research on Hindsight Bias. *Basic and Applied Social Psychology* **26**(2-3) 103–117.

Harley, Erin M. 2007. Hindsight Bias In Legal Decision Making. *Social Cognition* **25**(1) 48–63.

Hayward, Mathew L. A., Donald C Hambrick. 1997. Explaining the Premiums Paid for Large Acquisitions: Evidence of CEO Hubris. *Administrative Science Quarterly* **42**(1) 103.

Holt, Charles A, Susan K Laury. 2002. Risk Aversion and Incentive Effects. *American Economic Review* **92**(5) 1644–1655.

Howard, Michael. 1984. *The causes of wars and other essays*. Harvard University Press.

Huffman, David, Collin Raymond, Julia Shvets. 2019. Persistent Overconfidence and Biased Memory: Evidence from Managers. *Unpublished manuscript* (September 2017) 1–37.

John, Oliver P., Richard W. Robins. 1994. Accuracy and bias in self-perception: Individual differences in self-enhancement and the role of narcissism. *Journal of Personality and Social Psychology* **66**(1) 206–219.

Johnson, Dominic D. P. 2004. *Overconfidence and War: The Havoc and Glory of Positive Illusions*. Harvard University Press, Cambridge, MA.

Kennedy, Jane. 1995. Debiasing the curse of knowledge in audit judgment. *Accounting review: A quarterly journal of the American Accounting Association* **70**(2) 249–273.

Kessler, D., M. McClellan. 1996. Do Doctors Practice Defensive Medicine? *The Quarterly Journal of Economics* **111**(2) 353–390.

Kessler, Daniel P., Mark B. McClellan. 2002. How liability law affects medical productivity. *Journal of Health Economics* **21**(6) 931–955.

Keysar, B., A. S. Henly. 2002. Speakers' overestimation of their effectiveness monitoring and ambiguity. *Psychological Science* **13**(3) 207–212.

Koellinger, Philipp, Maria Minniti, Christian Schade. 2007. "I think I can, I think I can": Overconfidence and entrepreneurial behavior. *Journal of Economic Psychology* **28**(4) 502–527.

Kruger, Justin, Nicholas Epley, Jason Parker, Zhi-Wen Ng. 2005. Egocentrism over e-mail: Can we communicate as well as we think? *Journal of Personality and Social Psychology* **89**(6) 925–936.

Kőszegi, Botond. 2006. Ego Utility, Overconfidence, and Task Choice. *Journal of the European Economic Association* **4**(4) 673–707.

Loewenstein, George F, Don A. Moore, Roberto A. Weber. 2006. Misperceiving the value of information in predicting the performance of others. *Experimental Economics* **9**(3) 281–295.

Madarász, Kristóf. 2012. Information projection: Model and applications. *Review of Economic Studies* **79**(3) 961–985.

Malmendier, Ulrike, Geoffrey Tate. 2005. CEO Overconfidence and Corporate Investment. *The Journal of Finance* **60**(6) 2661–2700.

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

Malmendier, Ulrike, Geoffrey Tate. 2015. Behavioral CEOs: The Role of Managerial Overconfidence. *Journal of Economic Perspectives* **29**(4) 37–60.

Moore, Don A., Daylian M. Cain. 2007. Overconfidence and underconfidence: When and why people underestimate (and overestimate) the competition. *Organizational Behavior and Human Decision Processes* **103**(2) 197–213.

Moore, Don a., Paul J Healy. 2008. The trouble with overconfidence. *Psychological Review* **115**(2) 502–517.

Neale, M. A., M. H. Bazerman. 1985. The Effects of Framing and Negotiator Overconfidence on Bargaining Behaviors and Outcomes. *Academy of Management Journal* **28**(1) 34–49.

Newton, L. 1990. Overconfidence in the communication of intent: Heard and unheard melodies. *Unpublished doctoral dissertation, Stanford University, Stanford, CA* .

Niederle, Muriel, Lise Versterlund. 2007. Do Women Shy Away From Competition? Do Men Compete Too Much? *Quarterly Journal of Economics* **122**(3) 1067–1101.

Odean, Terrance. 1998. Volume, Volatility, Price, and Profit When All Traders Are Above Average. *The Journal of Finance* **53**(6) 1887–1934.

Rabin, Matthew. 1998. Psychology and Economics. *Journal of Economic Literature* **36** 11–46.

Rabin, Matthew, Joel L Schrag. 1999. First Impressions Matter: A Model of Confirmatory Bias. *The Quarterly Journal of Economics* **114**(1) 37–82.

Rammstedt, Beatrice, Oliver P John. 2007. Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality* **41**(1) 203–212.

Rensink, Ronald A., J. Kevin O'Regan, James J. Clark. 1997. To see or not to see: The Need for Attention to Perceive Changes in Scenes. *Psychological Science* **8**(5) 368–373.

Roese, Neal J., Kathleen D. Vohs. 2012. Hindsight Bias. *Perspectives on Psychological Science* **7**(5) 411–426.

Schaefer, Peter S., Cristina C. Williams, Adam S. Goodie, W.Keith Campbell. 2004. Overconfidence and the Big Five. *Journal of Research in Personality* **38**(5) 473–480.

Scheinkman, José A, Wei Xiong. 2003. Overconfidence and Speculative Bubbles. *Journal of Political Economy* **111**(6) 1183–1220.

Schwardmann, Peter, Joël van der Weele. 2019. Deception and self-deception. *Nature Human Behaviour* **3**(10) 1055–1061.

Simons, Daniel J, Daniel T. Levin. 1997. Change blindness. *Trends in Cognitive Sciences* **1**(7) 261–267.

Slovic, Paul, Baruch Fischhoff. 1977. On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance* **3**(4) 544–551.

Studdert, David M. 2005. Defensive Medicine Among High-Risk Specialist Physicians in a Volatile Malpractice Environment. *JAMA* **293**(21) 2609.

Thaler, Richard H. 2015. *Misbehaving: The Making of Behavioral Economics*. Business/Economics, W.W. Norton.

Wu, Brian, Anne Marie Knott. 2006. Entrepreneurial Risk and Market Entry. *Management Science* **52**(9) 1315–1330.

# Appendices

## A Additional Analysis

### A.1 Regressions with controls for individual absolute performance per work round

Table 3: Regressions of guessed tournament ranks on treatment and dummies for performance levels.

| Dependent variable: (Ordered probit) | Guessed tournament rank (1=best, 4=worst) (1) | (2) | (3) |
|---|---|---|---|
| Hindsight Treatment | $-1.036^{***}$ | $-1.071^{***}$ | $-1.011^{***}$ |
| | (0.197) | (0.195) | (0.193) |
| Dummies for absolute performance levels in | Round 1 | Round 2 | Round 3 |
| $N$ | 144 | 144 | 144 |
| $\log L$ | $-175.730$ | $-172.131$ | $-172.813$ |

Note: Estimated coefficients on performance dummies and cutoffs from ordered probit regressions are omitted. Values in parentheses are standard errors: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4: Regressions of piece-rate equivalents on treatment and dummies for performance levels.

| Dependent variable: (Tobit regressions) | Piece-rate Equivalent (1) | (2) | (3) |
|---|---|---|---|
| Hindsight treatment | $0.166^{**}$ | $0.180^{**}$ | $0.164^{**}$ |
| | (0.080) | (0.077) | (0.078) |
| Constant | $0.800^{***}$ | $1.320^{***}$ | $1.093^{***}$ |
| | (0.266) | (0.460) | (0.329) |
| Dummies for absolute performance levels in | Round 1 | Round 2 | Round 3 |
| $N$ | 144 | 144 | 144 |
| $\log L$ | $-96.417$ | $-94.267$ | $-96.933$ |

Note: Estimated coefficients on performance dummies are omitted. Values in parentheses are standard errors: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

## A.2 Hindsight-biased overentry by performance type

Table 5: Test of the treatment effect on tournament participation conditional on individual ability.

| Dependent variable: | Tournament Entry | | | | | |
|---|---|---|---|---|---|---|
| (Probit regressions) | Performance cutoff based on | | | | | |
| | Work Round 1 | | Work Round 2 | | Work Round 3 | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Hindsight treatment | 0.668* | 1.000** | 1.235** | 1.666** | 0.741* | 0.912** |
| | (0.384) | (0.454) | (0.518) | (0.652) | (0.415) | (0.452) |
| High Performer | −0.376 | −0.057 | −0.023 | 0.000 | −0.048 | 0.152 |
| | (0.391) | (0.448) | (0.372) | (0.420) | (0.365) | (0.419) |
| Hindsight×High Performer | −0.206 | −0.687 | −1.105* | −1.526* | −0.407 | −0.536 |
| | (0.590) | (0.699) | (0.651) | (0.824) | (0.573) | (0.665) |
| Risk lovingness (SOEP) | 0.134** | 0.101 | 0.145** | 0.126 | 0.130* | 0.098 |
| | (0.068) | (0.076) | (0.069) | (0.078) | (0.067) | (0.075) |
| Gender (1-male) | 0.540* | 0.532 | 0.496* | 0.522 | 0.549* | 0.548 |
| | (0.295) | (0.338) | (0.294) | (0.345) | (0.296) | (0.349) |
| High School GPA | 0.023 | 0.272 | 0.085 | 0.367 | 0.052 | 0.249 |
| | (0.250) | (0.308) | (0.261) | (0.334) | (0.249) | (0.304) |
| Age | 0.033 | 0.021 | 0.034 | 0.020 | 0.030 | 0.018 |
| | (0.035) | (0.037) | (0.035) | (0.037) | (0.033) | (0.034) |
| Extraversion (Big 5) | | 0.144 | | 0.309 | | 0.147 |
| | | (0.677) | | (0.714) | | (0.662) |
| Agreeableness (Big 5) | | 0.386 | | 1.073 | | 0.735 |
| | | (0.982) | | (1.046) | | (0.961) |
| Conscientiousness (Big 5) | | 0.444 | | 0.740 | | 0.513 |
| | | (0.786) | | (0.832) | | (0.790) |
| Neuroticism (Big 5) | | −2.651*** | | −2.391** | | −2.438** |
| | | (0.996) | | (0.997) | | (0.965) |
| Openness (Big 5) | | 0.815 | | 0.697 | | 0.692 |
| | | (0.661) | | (0.650) | | (0.647) |
| Generalized Optimism | | 0.040 | | 0.059 | | 0.039 |
| (Life Orientation Test) | | (0.040) | | (0.043) | | (0.040) |
| Constant | −0.754 | −0.870 | −1.046 | −2.073 | −0.832 | −1.097 |
| | (0.952) | (1.586) | (0.986) | (1.884) | (0.897) | (1.596) |
| $N$ | 144 | 144 | 144 | 144 | 144 | 144 |
| $\log L$ | −51.942 | −45.277 | −50.544 | −42.939 | −52.754 | −46.056 |

Note: Subjects are classified as low performers [high performers] if they solved fewer than [at least] six tasks in the round given in the header. Values in parentheses are standard errors: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

## A.3 Excluding participants who guessed they could solve all tasks

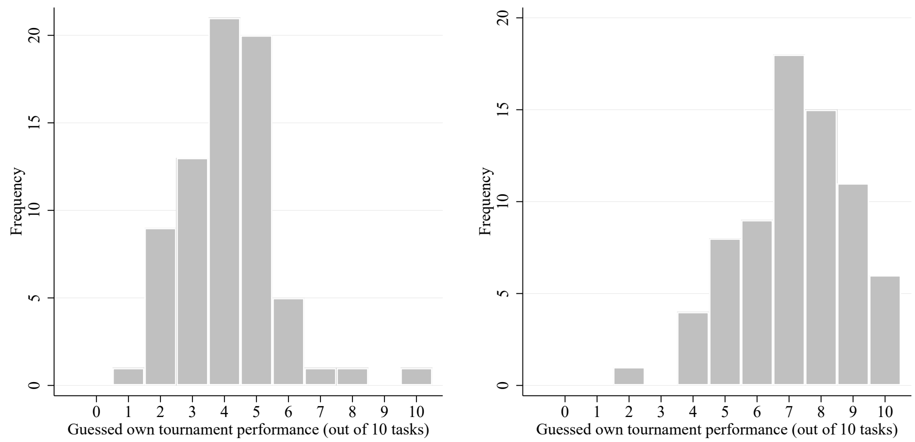


Figure 8: Distribution of guessed own absolute performance in the tournament after seeing the examples in the No-Hindsight treatment (left panel) and the Hindsight treatment (right panel).

Table 6: Regressions of guessed tournament ranks on treatment and performance dummies—excluding all participants who guessed that they could solve 10/10 tasks.

| Dependent variable: | Guessed tournament rank (1=best, 4=worst) | | |
|---|---|---|---|
| (Ordered probit) | (1) | (2) | (3) |
| Hindsight Treatment | $-1.048^{***}$ | $-1.075^{***}$ | $-1.027^{***}$ |
| | (0.203) | (0.200) | (0.197) |
| Dummies for absolute performance levels in | Round 1 | Round 2 | Round 3 |
| $N$ | 137 | 137 | 137 |
| $\log L$ | −167.551 | −164.778 | −166.376 |

Note: Estimated coefficients on performance dummies and cutoffs from ordered probit regressions are omitted. Values in parentheses are standard errors: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 7: Regressions of piece-rate equivalent on treatment, performance, and guessed tournament rank—excluding all participants who guessed that they could solve 10/10 tasts.

| Dependent variable: (Tobit regressions) | Piece-rate Equivalent (1) | (2) | (3) |
|---|---|---|---|
| Hindsight treatment | $0.175^{**}$ | $0.170^{**}$ | 0.021 |
| | (0.079) | (0.077) | (0.081) |
| Guessed rank in tournament | | | $-0.165^{***}$ |
| | | | (0.040) |
| Dummies for performance levels | No | Yes | Yes |
| Constant | $0.941^{***}$ | $1.330^{***}$ | $1.973^{***}$ |
| | (0.055) | (0.452) | (0.453) |
| $N$ | 137 | 137 | 137 |
| $\log L$ | −90.146 | −86.106 | −78.016 |

Note: Columns 2 and 3 include dummies for each performance level in the tournament (round 2). The results are qualitatively the same when adding dummies for each performance level in round 1 or round 3 instead of round 2. (Guessed) tournament ranks range from 1 (best) to 4. Values in parentheses are standard errors: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

# B Original screenshots (German)

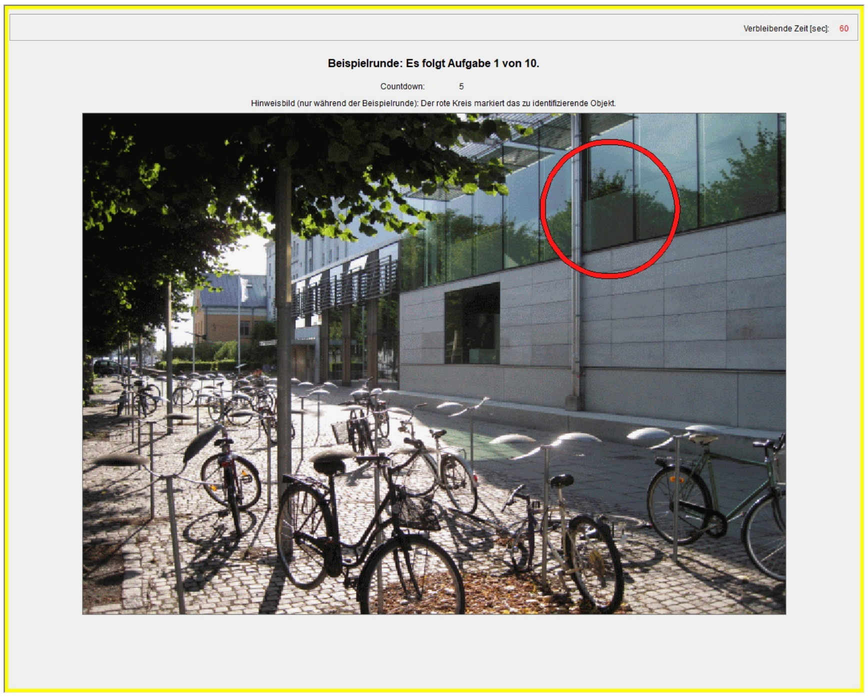


Figure 9: Solution to example task provided in the Hindsight treatment.

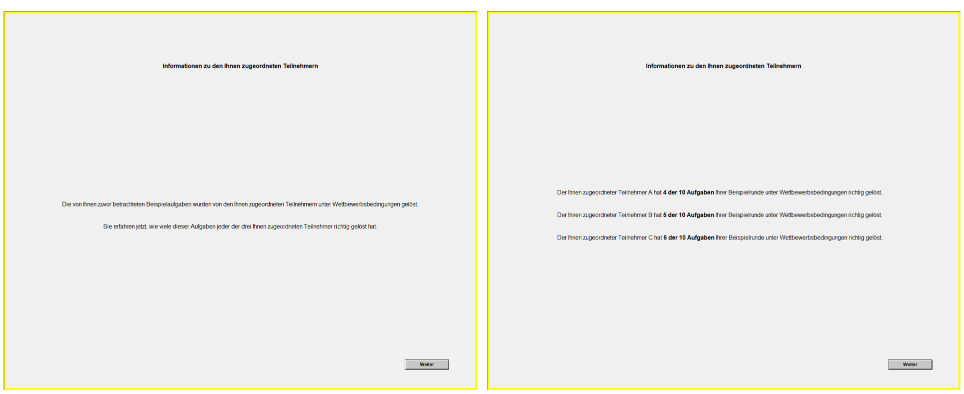


Figure 10: Information about matched participants. The reminder (left panel) was shown before the performance data (right panel).

# C Instructions for the informed and the uninformed treatments (translated from German)[34]

This experiment is part of a research project funded by the German Science Foundation. The experiment is on economic decision making.

**From now on, and for the entire time of the experiment, you are not allowed to talk to other participants or to communicate with them in any way. If you have a question, please raise your hand and we will come to you and answer your question. Please do not ask any question out loud. If you violate these rules, we will have to end the experiment.**

You will receive 5 Euro for participating in this experiment. You can earn additional money during the experiment. The instructions are the same for all participants.

### Overview

The experiment consists of several rounds. In each round, you will work on a fixed number of tasks. Your payment at the end of the experiment depends on your choices, on the number of tasks you solve correctly in one of the rounds, and, potentially, on the number of tasks that participants in previous experiments solved correctly. The exact timeline of the experiment, the decisions you have to make, and the payment procedures are explained in detail below.

### Structure of the Tasks

For each task, you will watch a short video clip, each consisting of two images shown alternately on the computer screen. The two images are identical except for one area. The difference between the two images will consist of one or more objects present in one of the images but partially or completely missing in the other image. Your task is to spot the difference between the two images.

Figure 1 shows an example. Image A in the example contains a kayaker on

[34] The treatment difference is indicated in `[square brackets]`.

the left side of the image. The kayaker is not present in image B.

Image A

Image B

Figure 1: Example of an image pair.

During the task, one image is shown for about one second. After a short blank screen, the other image is shown. The sequence of images is: Image A, Image B, Image A, Image B, . . . and so on. The total duration of each video clip is 14 seconds.

After the video clip, you have 40 seconds to submit an answer. During this time, you will see a response grid (see Figure 2). The response grid has numbered sections covering the image that contains the object to identify (Image A in the example).

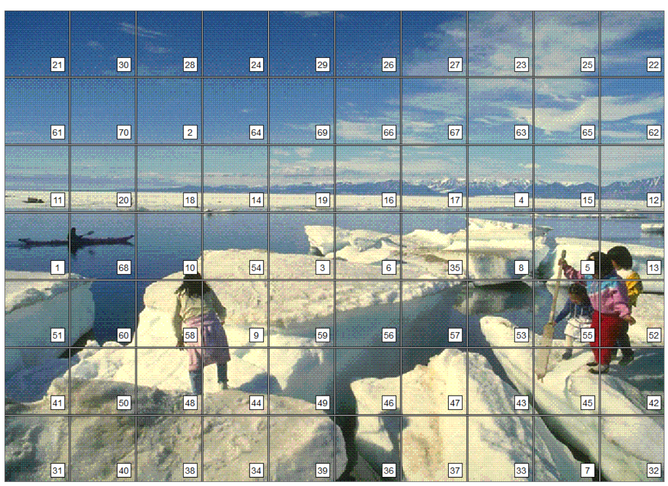

Figure 2: Example of response grid.

There will be an input field below the response grid. There, you can enter one of the grid numbers. To solve the task, you must enter the number of a section that contains the difference between the images. If the difference covers more than one section, the number of any section containing the difference is a valid answer. So, in the example above (Figure 2), you can enter “1” or “68” to solve the task correctly. *Please note that simultaneous entries of several grid numbers will not be evaluated.*

To proceed, click “Continue.” If you submitted an answer, it is now saved and shown for confirmation on the screen. You can now click “Next Task” to proceed or “Revise” to correct your input within the allowed response time.

*Please note that only your final answer will be evaluated.* Hence, if you make a correction—that is, you enter a new grid number in the input field and click “Continue” to save your answer—then only this and not the previous answer will be validated. If you do not complete a correction within the given response time, your previous answer will be evaluated.

The experiment will continue as soon as all participants have clicked “Next Task” or after the 40-second response time has elapsed.

## Matching to a Group with Participants in Previous Experiments

Participants in previous experiments have worked through all tasks you are going to see in the experiment.

*At the beginning of this experiment, three participants in previous experiments will be randomly matched to you.* You and your matched participants will form a group. This group assignment will remain the same for the entire time of the experiment; that is, the three participants matched to you will be the same for the entire duration of the experiment. The randomly selected participants of previous experiments are, however, not the same for all participants of this experiment.

To ensure anonymity, you will not learn the identity of the former participants with whom you are matched. Your anonymity will also be ensured during the entire experiment.

## Structure of the Rounds

Each round consists of ten tasks, that is, ten video clips, in which you have to spot the difference between two images. At the end of each round, your rank will be determined relative to your matched participants.

Your ranking is based on the number of tasks you solve correctly within a given round. You rank 1st if you solve more tasks than each of your three matched participants. If the number of tasks you solve is the second-highest, then you rank 2nd. If you solve the third-highest number of tasks, you rank 3rd. If you solve the lowest number of tasks compared to your matched participants, then you rank 4th.

If you and one or more of your matched participants solve the same number of tasks, then your ranking relative to these participants is determined randomly. For example, if you solve the same number of tasks as one of your matched participants,while the two other participants matched to solve fewer tasks than you, then you have the same chance to be ranked 1st as your matched participant who solves the same number of tasks as you.

Of course, the comparison of your score with those of your matched participants is based on exactly the same videos. That is, the number of your correctly solved tasks in a given round will be compared to the number of correctly solved tasks of your matched participants for the exact same videos.

## Conditions for Your Matched Participants

Your matched participants previously performed the tasks under the *same conditions as you.* Specifically, they had the same time as you to work on the tasks: 14 seconds to watch the video and 40 seconds to submit an answer.

However, your matched participants worked on all the tasks that you will see in this experiment under a competitive payment scheme—including the tasks in your Example Round. Specifically, they received a payoff in a given a round only if they ranked 1st within their group.

## Timeline and Payment

The timeline of the experiment is as follows:

1. Decision Phase 1
2. Example Round (10 tasks)
3. Information about your matched participants
4. Decision Phase 2: Choice of Your Critical Piece Rate
5. Announcement of the Piece Rate in the envelope
6. Round 1 (10 tasks): Piece Rate
7. Round 2 (10 tasks): Competition
8. Round 3 (10 tasks): Piece Rate or Competition
9. Questionnaire
10. Random draws and calculation of your payment
11. Feedback and payment

## 1 Decision Phase 1

At the beginning of the experiment, you will face a number of decision-making problems. At the end of the experiment, one of your decisions will be randomly selected for payment. You will receive more information about the decision-making problems and the payment rules during the experiment.

## 2 Example Round

After Decision Phase 1, you will participate in an Example Round. Here, you will see ten tasks in the same way as in the subsequent three rounds. You cannot earn money in the Example Round; that is, your input will not be evaluated. `[Hindsight Treatment only:` During the Example Round, you receive a guide to the solution at the outset of each task. You will not receive solution guides during the following payoff-relevant rounds.`]`

Both the allocation of the videos to your Example Round and the three subsequent rounds as well as order of the videos within each round, are the result of a random draw. This random draw does not take into account the order in which your matched participants solved the tasks. The order of the tasks was also determined randomly for these participants.

Because the tasks are not precisely the same in terms of their degree of difficulty, the random draw for this experiment ensured a uniform degree of difficulty across Rounds. That is, both the Example Round and the following three Rounds contain the same ratio of comparatively easier and harder tasks.

The random draw of the videos is the same for all participants of this experiment; that is, at a given time in this experiment, each participant will be watching the same video clip.

## 3 Information Phase

At this time, you will receive information on how your three matched participants did in your Example Tasks: For each of your three matched participants, you will learn how many of your Example Tasks they solved correctly. *Please note that your matched participants previously worked on your Example Tasks (as well as the tasks in Rounds 1–3) under Competition.* That is, they were paid for a given round only if they ranked 1st within their group.

## 4.–8. Decision Phase 2, Announcement of the Piece Rate, and Rounds 1–3

You will receive 7 Euro for working on the tasks in Rounds 1 to 3. *In addition, one of the Rounds 1–3 will be randomly selected at the end of the experiment.* There are two different payment conditions in the three Rounds.

### Round 1: Piece Rate

In preparation for this experiment, a Piece Rate was determined, sealed in an envelope, and attached to the door of the laboratory. The Piece Rate is between 0.01 Euro and 2 Euro. The envelope will be opened *after Decision Phase 2*, that is, before Round 1, and the Piece Rate will be entered in one of your computers.

If Round 1 is selected for payment, you will receive the Piece Rate in the envelope for each task that you solve correctly in this round.

### Round 2: Competition

If Round 2 is selected for payment, your payoff will be determined as follows: You receive 2 Euro for each correctly solved task if you rank 1st in this round;

otherwise, you will receive 0 Euro.

### Round 3: Piece Rate or Competition

If Round 3 is selected for payment, your payoff depends on your choice in *Decision Phase 2*, where you choose *Your Critical Piece Rate*.

Your choice of Your Critical Piece Rate determines how your payoff is calculated in Round 3 as follows:

- If the Piece Rate in the envelope is *smaller than* Your Critical Piece Rate, then you will be paid according to the Competition condition. That is, you will receive 2 Euro for each task you solve correctly in this round if you rank 1st in this round; otherwise, you will receive 0 Euro.
- If the Piece Rate in the envelope is *greater than or equal to* Your Critical Piece Rate, then you will be paid according to the Piece Rate in the envelope. That is, you receive the Piece Rate in the envelope for each task you solve correctly in this round.

During the experiment, a graphical interface will assist you in your decision making. For each participant, the starting value of the Critical Piece Rate in the graphical interface is randomly chosen by the computer.

### Questions During the Experiment

You will be asked a total of nine questions at different times during the experiment. Here, you can earn additional money. At the end of the experiment, one of the nine questions will be randomly selected. You will receive 1 Euro if you answered this payoff-relevant question correctly. You will receive instructions for the exact payment rules for the questions during the experiment.

## 9 Questionnaire

After the three rounds, you will receive a questionnaire to complete. Although your answers will not affect your payoff, we ask that you answer the questions carefully.

## 10.–11. Random Draws, Feedback, and Payment

At the end of the experiment, one of your choices from Decision Phase 1, one of your answers to the nine questions during the experiment, and the payment round (the same for all participants) will be randomly selected for payment. Your respective payoffs and your total payment will be shown on the screen. You will also receive information about all three rounds—in particular, the number of tasks you solved correctly in each round and your rank in each round.

Your total payment for this experiment, thus, consists of:

1. Your show-up fee (5 Euro);
2. Your payment for working on the tasks (7 Euro);
3. Your payoff according to one of your choices in Decision Phase 1;
4. Your payoff according one of your answers to the questions during the experiment; and
5. Your payoff according to the randomly selected payoff-relevant round.

Do you have questions? If yes, please raise your hand. We will answer your questions in private.

**Thank you for participating in this experiment!**

## Comprehension Questions

1. After the Example Round, you will receive information about your matched participants. Who are these participants?
2. Your score in Rounds 1–3 will be used to calculate your rank relative to other participants. Who are these participants?
3. Do questions 1 and 2 refer to the same participants?
4. At what time will you be matched to a group; that is, at what point in the experiment will other participants be matched to you?
5. How many participants will be matched to you?
6. Which of the tasks that you will see in your Example Round and Rounds 1–3 of this experiment were also performed by your matched participants?
7. Under which payment condition did your matched participants perform the tasks of Round 1–3?
8. Under which payment condition did your matched participants perform the tasks of your Example Round?
9. The Piece Rate is relevant for which rounds?
10. What is the amount of the Piece Rate, and when do you learn the amount of the Piece Rate?
11. How is your payoff calculated if Round 1 is randomly selected for payment?
12. How is your payoff calculated if Round 2 is randomly selected for payment?
13. Imagine that Round 3 is randomly selected for payment. Assume, further, that in Decision Phase 2, you set Your Critical Piece Rate to X Euro and that the Piece Rate in the envelope is Y Euro.
    (a) How is your payoff in Round 3 calculated if $X > Y$, that is, if Your Critical Piece Rate is larger than the Piece Rate in the envelope?
    (b) How is your payoff in Round 3 calculated if $X \leq Y$, that is, if Your Critical Piece Rate is smaller than or equal to the Piece Rate in the envelope?