The Influence of Overconfidence and Competition Neglect On Entry Into Competition

Katharina Schüssler (LMU Munich)

Discussion Paper No. 87

March 26, 2018

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Dr. Klaus M. Schmidt, University of Munich, 80539 Munich, Germany
+49 (89) 2180 3405 | info@rationality-and-competition.de

The Influence of Overconfidence and Competition Neglect on Entry into Competition*

Katharina Schüßler†

March 14, 2018

I investigate whether two mechanisms leading to biased beliefs about success, overconfidence and competition neglect, influence decisions to enter competitive environments. I use a controlled laboratory setting that allows to elicit belief distributions related to absolute as well as relative overconfidence to study it comprehensively and introduce two treatment variations: First, some participants receive detailed performance feedback addressing absolute and relative overconfidence before making their decision. Second, I vary whether the competition group consists of all potential competitors or only of individuals who also chose to compete. I find that there is systematic heterogeneity in perception biases. In addition, both mechanisms influence individuals' decisions. However, choices are closely tied to previous performance and assessments, and there are no significant gender differences.

Keywords: Competition neglect; competitive behavior; feedback; overconfidence

JEL Classification Numbers: C91; D83; J16

* The author is grateful to Maximilian Breu, Florian Englmaier, Matthias Fahn, Nadine Geiger, Raphael Guber, Stephen Leider, Yves Le Yaouanq, Simeon Schudy, Michael Schüßler, Peter Schwardmann, and participants of the Experimental Brown Bag seminar at LMU Munich for valuable comments and suggestions. The author gratefully acknowledges funding through the International Doctoral Program "Evidence-Based Economics" of the Elite Network of Bavaria and and financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190.

† Department of Economics, University of Munich. Address: Kaulbachstr. 45, D-80539 Munich, Germany. E-mail: *katharina.schuessler@econ.lmu.de*.

1. Motivation

Traditionally, in contract theory, the rationale for providing incentives is to align the interests of principals and agents in order to overcome contracting problems in the presence of information asymmetries (Hölmstrom, 1979; Grossman and Hart, 1983). However, this view underestimates the importance of worker self-selection, i.e., that different types of workers are attracted by different kinds of incentive schemes and therefore systematically select into different environments. When considering both aspects, performance does not only depend on an incentive effect, but also on a sorting effect (e.g., Lazear, 2000; Dohmen and Falk, 2011; Larkin and Leider, 2012).

Selection into competitive environments has proven to be an important factor in many economic settings like labor markets (based on rank-order tournaments; cf. Dechenaux et al., 2015) or entrepreneurship (Camerer and Lovallo, 1999). Findings reveal that factors other than productivity influence selection into competitive environments (e.g., Gneezy et al., 2003; Niederle and Vesterlund, 2007; Vandegrift et al., 2007; Bartling et al., 2009; Eriksson et al., 2009; Dohmen and Falk, 2011; Balafoutas et al., 2012). Papers investigating determinants of entry decisions from different angles have identified several economic preferences that influence entry into competitive environments: risk aversion and distributional preferences matter as well as heterogeneity in attitudes towards competition. Aside from preferences, differences in overconfidence have also been brought forward as explanatory mechanism. However, in this area, central blind spots remain as overconfidence comprises different facets that need to be carefully measured, and alternative mechanisms have not been systematically explored.

First, rather than being a clearly defined mechanism, overconfidence is an umbrella term containing several different biases in beliefs or, more precisely, errors in performance assessments, that have been discussed controversely. Two different facets of overconfidence identified by Moore and Healy (2008), the overestimation of one's ability (overestimation) and the belief to be better than others (overplacement), might play a role for entry decisions. Different aspects have been emphasized by different literature streams: while some have focused on overestimation of own ability (the literature related to Camerer and Lovallo, 1999), others concentrate on relative overconfidence or overplacement while effectively eliminating uncertainty about own performance (the literature buliding on Niederle and Vesterlund, 2007), with comprehensive approaches being scarce.[1]

Second, overconfidence in beliefs has mostly been described as being caused by subjective priors or updating biases like attaching excessive precision to signals (Jehiel, 2016). However, other mechanisms that lead to biased perceptions might play a role as well. Specifically, it has been argued in research on bounded rationality that individuals do not realize that others' actions depend on their private information. This notion is incorporated in solution concepts like cursedness (Eyster and Rabin, 2005) or behavioral equilibrium (Esponda, 2008) and has been utilized as driver of biased beliefs using different terminologies – for instance, Enke and Zimmermann (2017) show that correlation neglect leads to biased information processing, and Jehiel (2016) derives overoptimism from selection neglect. Closest to the setting of this paper, Camerer and Lovallo (1999) identify competition neglect, that is, that individuals tend to neglect self-selection of (potentially high-performing) others into competition as a potential mechanism biasing entry decisions. However, since this mechanism has been brought up as a potential cause for excess entry, only a handful of papers have followed up on it. On the contrary, most studies rule self-selection of competitors out by design, and systematic identification of the role of this bias for competition entry remains sparse.

This study aims at investigating whether overconfidence and competition neglect influence decisions to enter competitive environments. I use a controlled laboratory setting that allows to elicit belief distributions related to absolute as well as relative overconfidence to study it comprehensively while influencing it via performance feedback. In addition, my setting allows me to introduce systematic variation in the possibility for competition neglect. Studying these two mechanisms simultaneously is relevant as both overconfidence and competition neglect potentially lead to the same behavioral pattern, namely excess entry. However, for de-biasing, it is essential to know whether individu-

[1] The experiment by Kamas and Preston (2012) constitutes a notable exception as the authors explicitly relate to the proposed systematization by Moore and Healy (2008).

als overestimate their chances of success because they overestimate their performance in absolute terms or relative to others, or because they do not realize that they face a selected set of competitors.

With this in mind, I let individuals decide whether they want to enter a competition in a real-effort experiment while eliciting their beliefs with respect to their own and others' performance as well as their winning probability. To answer my research question, I introduce two treatment variations: First, some participants receive detailed performance feedback addressing absolute and relative overconfidence before making their decision. Comparing entry decisions of individuals who have received feedback to those who have not enables me to investigate whether overconfidence plays a role for selection into competition – if there is a difference between those groups, it does. Second, I vary whether the competition group consists of all potential competitors or only of individuals who also chose to compete. Here, finding no difference in entry between the groups implies that decision makers fail to adjust their decisions to the selected sample, thus exhibiting competition neglect.

The main findings of this paper are as follows: Individuals exhibit pronounced heterogeneity with respect to their performance beliefs, with low-performing individuals overestimating their own performance and their chances of success while underestimating performance in the competition group, and the opposite being true for high-performing individuals. These biases in performance beliefs are ameliorated by feedback, but the composition of the competition group is persistently disregarded. Investigating determinants of entry decisions to tackle the key question of this paper whether competition neglect and overconfidence influence entry into competition, I find that both influence individuals' decisions. However, competition entry closely tied to previous performance and assessments, and in contrast to the literature, there are no gender differences in beliefs and entry decisions.

The remainder of this chapter is organized as follows. Section 2 gives an overview of related literature. Section 3 presents a simple theoretical framework illustrating the main argument of this paper that overconfidence and competition neglect lead to the same behavioral changes while constituting different mechanisms. This is the prerequisite of the experimental design, which is presented in Section 4. Section 5 describes and analyzes the experimental results. Section 6 discusses and concludes.

2. Related Literature

This study relates to several strands of literature. First, there is a wide variety of research focusing on overentry into markets of entrepreneurs, which are highly competitive environments (see, e.g., Camerer and Lovallo, 1999, for an early lab experiment). In this setting, empirical evidence suggests that an entrepreneur's confidence in his own skill, ability and knowledge has a crucial impact on new business creation and might explain the high failure rate of new business owners (Koellinger et al., 2007; Cain et al., 2015; Bolger et al., 2008). Apart from overconfidence, the neglect of selective entry of competitors into a market may be a leading cause of failure: Even if an entrepreneur assesses his skills correctly, he may still be not aware of the fact that his competitors are not a random sample of the population, but have abilities that are at least as good as his. The competitors' skills may be even higher, since their business has survived years of constant pressure from market powers that drive out non-competitive participants (Camerer and Lovallo, 1999; Moore et al., 2007). Competition neglect has been shown to also matter for market entry decisions in other situations like selling on eBay (Simonsohn, 2010). However, systematic identification of competition neglect remains scarce. Thus, my study contributes to this strand of literature by experimentally varying the possibility of competition neglect.

Second, entry into competition has been extensively studied from a different angle: over the course of the last decade, a large number of studies has explored gender differences in competition decisions (see Niederle and Vesterlund, 2011, for an overview). In particular, one worrying finding is that high-performing women appear to opt out of competitive environments (Niederle and Vesterlund, 2011). Building on the seminal work by Niederle and Vesterlund (2007), who focus on individual-level competition decisions and find a large gender gap, it has been suggested that women show a lower willingness to engage in competition, a finding that has also been replicated in real-world occupation settings (Flory et al., 2015). Regarding the mechanisms driving this, a plethora of laboratory evidence has been produced. At first, differences were mostly attributed to heterogeneity in attitudes towards competition as well as risk and feedback aversion (Eckel and Grossman, 2008). Lately, it has been argued, however, that differences in beliefs could explain the above-described outcomes. Psychology research suggests that women exhibit less overconfidence than men, a finding that has been recently incorporated in economic investigations (see Niederle and Vesterlund, 2011). My study thus contributes to the literature on gender and competition by investigating more closely how biased beliefs about the performances of oneself and the competition contribute to differences in selection into competition.

Third, my study is related to the literature addressing the potentially benefiting role of feedback in reducing overconfidence effects (see Arkes et al., 1987, for an early paper on this). Several papers study how subjects update their relative performance beliefs and the relation to competition decisions. A general finding is that while there is belief updating, after noisy feedback, individuals tend to update less than a Bayesian would (Möbius et al., 2014). For instance, Buser et al. (2016) show that there is heterogeneity in this conservatism, with women being more conservative than men. Wozniak et al. (2014) find that providing relative feedback causes more high-performing women and less low-performing men to select into competition, thus reducing the gender gap in competition decisions. Berlin and Dargnies (2016) use a setting close to mine (except for only giving binary feedback) and show that different aspects of performance matter for men and women. In contrast to most of this literature, I give full performance feedback. In this setup, I also see sensible, albeit conservative belief adjustment.

Before going into details of the experiment, I now describe a simple theoretical framework of how overconfidence and competition neglect influence decision making.

3. Theoretical Framework

The goal of this section is to stress the central argument of this paper that multiple causes of biased beliefs leading to biased entry into competition exist that need to be addressed in different ways: overconfidence and competition neglect. Thus, the model illustrates how the two mechanisms lead to the same changes in behavior albeit being conceptually different.

3.1. Model Setup

I assume that an individual i can be of two types, $\theta_i \in \{\theta_L; \theta_H\}$, where $\theta_H > \theta_L$. θ_i represents the productivity of the individual in some task. There are N individuals who all face an identical game with the following structure that is common knowledge: First, nature draws the type of all players $i \in \{1, .., N\}$. Here, type θ_H is assigned with probability p. Second, every agent receives a private signal $S = \theta_i + \epsilon$ with $\epsilon \sim \mathrm{N}(0, \sigma^2)$, which determines the posterior probability q_i the individual assigns to being of type θ_H.[2] Hence, there is individual heterogeneity with respect to posterior beliefs q_i. Third, every individual is facing a choice between two options: to enter a competition or to receive an individual payment. If competition is chosen, the individual is matched with a randomly drawn opponent j who has also chosen to compete.[3] She receives a piece rate of p_C if her productivity is higher than that of her opponent and nothing otherwise. If both individuals are of the same type, the winner is randomly chosen. If the individual payment is chosen, a lower piece rate of $p_{NC} < p_C$ is received for sure.

3.2. Rational Benchmark

Assuming risk neutrality, individual i's expected utility from choosing competition equals

$$\mathbf{E}_C = p_C \times [q_i\theta_H \times (0.5q_j + (1 - q_j)) + (1 - q_i)\theta_L \times 0.5\,(1 - q_j)],$$

where $q_j = \Pr(\theta_j = \theta_H |\, j \text{ enters})$. In this case, her expected utility equals

$$\mathbf{E}_{NC} = p_{NC} \times [q_i\theta_H + (1 - q_i)\theta_L].$$

[2] This structure was chosen to generate continuous posterior beliefs about being of type θ_H. This has the advantage that it allows to identify threshold equilibria and analyze comparative statics.

[3] Note that I assume here that there is always at least one competitor as allowing for the possibility that there is none would make the problem more complicated while not affecting the main conclusions. Furthermore, this case did never occur in the experiment.

The individual decision is made by comparison of the expected utilities in both cases: if the expected utility from entering the competition exceeds the expected payoff from the individual payment, the individual enters. Then, she is indifferent between entering the competition and receiving an individual payoff at

$$q_i^*(s, \xi, q_j) = \frac{2-(1-q_j)s}{\xi[s(2-q_j)-2]-[(1-q_j)s-2]}$$

where $s = \frac{p_C}{p_{NC}} > 1$ and $\xi = \frac{\theta_H}{\theta_L} > 1$.

I focus on the symmetric Bayesian Nash equilibrium, which is characterized as follows:

Proposition 3.1. *For any q_i and $q^* = \frac{2-(1-q_j)s}{\xi[s(2-q_j)-2]-[(1-q_j)s-2]}$, there exists an equilibrium in pure threshold strategies where $\forall i \in \{1, ..., N\}$, individual i enters the competition if $q_i > q^*$, and chooses the individual payoff if $q_i < q^*$.*

To establish $q_i \lesseqgtr q^*$ as sensible decision criterion, q^* has to exist, be monotone and unique. Proofs for existence and uniqueness of q^* are straightforward. For monotonicity, I investigate the effects of s, ξ, and q_j on q^* and show that all of them are unambiguous.[4] First, the comparative static of q^* with respect to the piece rate ratio s is

$$\frac{\partial q^*}{\partial s} < 0.$$

In terms of the model, this means that as the piece rate ratio increases, the cutoff for entering the competition decreases. Intuitively, this makes sense as the piece rate ratio is a measure for the relative attractiveness of the competition that increases in the difference between the competitive and individual piece rates. Hence, as competition becomes more attractive compared to the individual incentives, the cutoff value decreases and more individuals enter the competition. Second, the comparative static of q^* with respect to the productivity ratio ξ is

$$\frac{\partial q^*}{\partial \xi} < 0.$$

[4] The proofs for existence and uniqueness as well as the derivatives and the determination of their signs can be found in Appendix A.

This means that as the performance difference between the high type and the low type increases, the decision cutoff decreases, which makes intuitive sense: as relative productivity associated with being of type θ_H increases, expected utility from choosing competition rises faster than expected utility from choosing the individual pay as the piece rate is higher for the former than for the latter. Third, the comparative static of q^* with respect to the probability that the competitor is of type θ_H, q_j, is

$$\frac{\partial q^*}{\partial q_j} > 0.$$

Thus, the decision cutoff increases in probability that the competitor is of type θ_H. This result is intuitive as well as an increase in this probability makes it less likely that the competition is won. This way, an increase in q_j decreases the expected utility from competition while it does not affect expected utility from individual payment.

3.3. Introducing Biased Beliefs

Having established the cutoff value q^* as an equilibrium, I now introduce overconfidence and competition neglect and derive predictions how the cutoff and thus the fraction of individuals entering competition changes. Throughout, I assume full naiveté about the biases; that is, all players assume that they themselves and all other players act according to the rational benchmark.[5]

For overconfidence, I follow Moore and Healy (2008) and focus on two different aspects, overestimation of own performance and underestimation of others' performance, as both are potential ways to generate overplacement.

First, to introduce a simple notion of overconfidence, assume that the true probability of being of type θ_H is q_i^{true}. Rather than having realistic beliefs about own ability, i.e., $q_i = q_i^{true}$, an individual subject to overconfidence believes that $q_i = \hat{q}_i$ with $\hat{q}_i > q_i^{true}$. Then, beliefs that satisfy $q_i^{true} < q^* < \hat{q}_i$ lead to competition entry even though this is not payoff-maximizing. This leads to the following prediction:

[5] Of course, this is a very strong assumption that is made to keep the analysis as simple as possible.

Prediction 1. *Overestimation of own performance leads to overentry into competition.*

An additional way to introduce overconfidence is via the belief about the type of the competitor, q_j. More specifically, assume that an overconfident agent underestimates the competition, i.e. $\overline{q}_j < q_j^{true}$. As already discussed above, the cutoff value q^* is positively related to q_j. Thus, if the agent underestimates the probability of facing a strong competitor, she is also going to underestimate the cutoff value: $q^*(\overline{q}_j) < q^*(q_j^{true})$. This leads to the following prediction:

Prediction 2. *Underestimation of competitor's performance leads to overentry into competition.*

Second, to investigate the effect of competition neglect on competition choices, the belief about the probability that the random competitor is of type θ_H, q_j, is central. Remember that the cutoff value for selecting into competition increases in q_j. To understand how competition neglect influences this value, note that q_j is a conditional probability as the competitor has also made the choice to compete. Assuming rational agents, $q_j^{true} = \Pr(\theta_j = \theta_H | \; j \text{ enters}) \in [q^*, 1]$. In contrast, if the individual is subject to (full) competition neglect, she does not take into account the choice made by the competitors and considers the unconditional probability that her competitor is of type θ_H, $\hat{q}_j = \Pr(\theta_j = \theta_H) \in [0; 1]$. It follows that $\hat{q}_j < q_j^{true}$, which in turn leads to a decrease in the cutoff value: $q^*(\hat{q}_j) < q^*(q_j^{true})$.

Prediction 3. *Competition neglect leads to overentry into competition.*

This simple model sketch illustrates the argument central to this paper, namely that overconfidence and competition neglect lead to similar changes in behavior compared to the rational benchmark – both lead to excess entry into competition. Thus, they cannot be disentangled by only looking at behavioral outcomes. As a consequence, both potential mechanisms have to be investigated carefully in an experiment exploring causes of biased entry. The following section presents the experimental design utilized in this study in detail. It is designed to help explore the role of overconfidence and competition neglect for competition decisions.

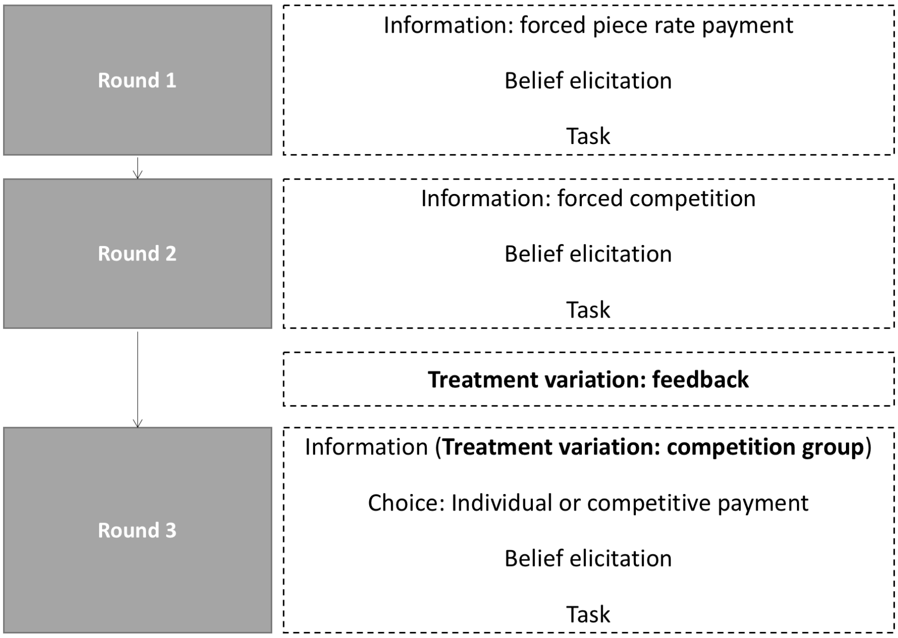
4. Experimental Design

The experiment proceeded as follows: Upon arriving at the laboratory and being randomly assigned to cubicles, participants were given general instructions and told that they would work on two unrelated parts. They were informed that their payment would depend on their performance and actions in the experiment, but that only one of the rounds or choices in each of the two parts would be paid out. To ensure incentive compatibility, participants would only be notified of the payoff-relevant rounds, decisions and performance at the very end of the experiment. Then, the main part of the experiment started immediately. In each of several rounds, participants had to work on different tasks: They performed a real-effort task under varying incentive schemes and had to answer several belief elicitation questions. In the last of three rounds, participants also had to make a choice between individual and competitive pay. Beliefs were elicited in every round. To experimentally vary overconfidence and competition neglect, I included two types of between-subject variation: depending on treatment, participants either received performance feedback or not, and either faced a random competitor or one that chose to compete. A summary of the procedure of this part is shown in Figure 1, and details about the different elements of the main part of the experiment are provided in the following subsections.[6] After completing this part, risk preferences of participants were elicited as these have been shown to be an important determinant of selecting into competitive environments. This was done using the standard choice list procedure following Holt and Laury (2002). After this second part was completed, participants were informed about their performance and payoffs and had to answer some general socio-economic questions.[7]

[6] Note that the experiment also incorporated a fourth round, which required subjects to choose between incentive schemes, but based on past rather than future performance. This round was conducted for robustness only and is thus not included in the main part of this paper. The results are qualitatively similar for those of round three. For the interested reader, a short description of this round and results are provided in Appendix D.

[7] These included gender, age, field of study, math grade as a crude proxy for general mental ability, familiarity with experiments, and mood.

Figure 1: *Experimental procedure in the main part of the experiment*



4.1. Real-Effort Task

In the real-effort task, participants had to correctly add as many blocks of five two-digit numbers as possible in three minutes. This task is commonly used in the experimental literature because it is easy to explain and produces substantial variation in individual performance (Niederle and Vesterlund, 2007). Subjects were not allowed to use a calculator, but were given a pen and scratch paper that they were welcome to use. In each round of the real-effort task, five randomly selected two-digit numbers were depicted on the screen in a row, with an additional field for the solution. Upon typing in their solution, participants had to confirm by clicking a "Submit" button and were then given the next five numbers. In contrast to Niederle and Vesterlund (2007), subjects were not forced to try one problem until they got the answer right, and they were not informed about the number of problems they had already solved correctly in that round. This was mainly done because this way, there is a greater scope for overestimation of own performance.[8] In addition to staying close to the related literature, I chose this

[8] In addition, the authors also state that they chose to force participants to answer a problem correctly before being able to proceed to prohibit them to skip hard problems and search for easy ones. In my case, there was no subject showing this behavior.

task because it can be seen as ego-relevant for a student subject pool and is therefore more likely to produce overly favorable judgments (Grossman and Owens, 2012) and because it has been argued in the literature there should not be a significant gender gap in performance in this task (e.g., Niederle and Vesterlund, 2007; Buser et al., 2014).

4.2. Incentives and Reference Group

After being told that they would have to work on the real-effort task and being given one minute to practice, participants faced three rounds with the same basic setup: In the beginning of each round, participants were informed about the payment rules. Then, performance beliefs were elicited before the actual task was conducted.

The rounds mainly differed in incentives. More specifically, following Niederle and Vesterlund (2007), in the first round, all subjects received a piece rate of 0.50 € for every correctly solved item. In the second round, all subjects then faced a competitive pay: they received a higher piece rate of 1.00 € per correctly solved item, but only if their performance in that round was higher than that of a randomly chosen opponent from the reference group. For round three, participants then could choose between the lower piece rate of 0.50 € and the competitive payment scheme.

Here, it is important to note that the reference group for this and every other competitive incentive scheme as well as for the corresponding belief elicitations explained in the following subsection did not consist of the other participants of the same experimental session, but of participants of a past session. This was mainly done for two reasons: On the one hand, it ensured that it was possible to inform participants about choices of the reference group before they had to make their own choices, and on the other hand, this way every observation in one session can be treated as an independent observation, which substantively increases the statistical power of the analyses. In addition, individual behavior could not affect others' payoff, ruling out any influence of social preferences on competition choices.

4.3. Treatment Variations

To get at overconfidence and competition neglect, two treatment variations were implemented in the main part of the experiment: First, to address overconfidence, feedback on own and others' performance was given between rounds two and three in half of the sessions. More specifically, participants in the feedback condition were told the number of items they had solved in round two as well as the average number of items solved and the whole performance distribution in the reference group. Second, the possibility for competition neglect was manipulated by design as in half of the sessions, the competitive pay was dependent on the performance of a randomly drawn participant of the reference group, while in the other half, the competition group in rounds three and four only consisted of participants of the reference group who had also chosen to work under the competitive incentive scheme in that round. This way, participants in the selection condition also had to factor in that they were facing a selected sample of better-performing individuals. Table 1 gives an overview of the treatment conditions.

Taken together, these treatment variations make it possible to explore how feedback about own and others' performance affects participants' beliefs and choices, if and how strongly participants are able to factor in the selectiveness of their competitors, and how this can be influenced by feedback as well.

Table 1: *Overview of the treatment conditions* [9]

		Competition Group	
		Selected	Random
Performance Feedback	No	None	Selection
	Yes	Overconfidence	Both

[9] Treatment conditions are named according to the addressed mechanism. Hence, in *Selection*, competition neglect is eliminated while overconfidence is not addressed, and in *Overconfidence*, overconfidence is tackled by feedback while the possibility for competition neglect persists.

4.4. Belief Elicitation

Because belief elicitation is crucial for addressing my research questions, beliefs are elicited in every round rather than at the very end of the experiment. In addition, instead of only eliciting participants' subjective ranks, I ask a total of three questions to get at the different facets of errors in beliefs and, in this way, at overconfidence and competition neglect.

First, it is important to keep in mind that concerns have been voiced that "overconfident" behavior is consistent with Bayesian updating in many cases. According to Benoît and Dubra (2011), this can be tracked back to the fact that the conventionally used research methodology does not allow to disentangle true overconfidence from rational information processing. Merkle and Weber (2011) identify aggregation of beliefs as greatest concern and show that eliciting belief distributions allows to rule out rational belief updating as explanation. I address these concerns by eliciting the whole belief distribution for own performance in the task as well as performance of a randomly drawn individual of the reference group. To do so, I implemented the procedure proposed by Harrison et al. (2015, 2017). More specifically, beliefs were elicited using a range of possible responses that were presented as ten intervals. Subjects then could allocate a total of 100 tokens according to their subjective beliefs. The allocation translated into payoffs according to the quadratic scoring rule with a maximum payoff of 2 € per question. Harrison et al. (2017) show that using this approach allows to elicit distributions while keeping the bias caused by risk aversion minimal.[10] Thus, I use these two questions to assess absolute and relative overconfidence reflected in performance beliefs.

Second, I elicited participants' subjective probability of performing better than a randomly chosen participant of the reference group using a reservation probability or crossover mechanism as first proposed by Ducharme and Donnell (1973) and implemented in variants by Möbius et al. (2014), Buser et al. (2016) or van Veldhuizen (2017), among others.[11] In particular, subjects had to indicate the probability p_s that

[10] Subjects were given detailed instructions explaining that the allocation of tokens to intervals comprised a trade-off between payoffs and risk. The full instructions can be found in Appendix E.

[11] See Schlag et al. (2015) for an overview of this and other belief elicitation mechanisms.

made them indifferent between receiving either a price of 2 € if their performance is better than that of a randomly drawn opponent or 2 € with that probability. Then, a random number z was drawn from a uniform distribution on the interval $[0, 1]$ and subjects received 2 € with probability z if $z > p_s$ and 2 € if their performance was better than that of a randomly drawn opponent otherwise. This procedure makes stating the true probability incentive compatible.[12] This question functions as compound measure for overconfidence and competition neglect.

5. Results

The experiment was conducted at MELESSA, the experimental laboratory of the University of Munich, in June of 2017. A total number of 240 participants were recruited using Orsee (Greiner, 2015) and took part in 10 sessions (two for deriving the performance of the reference group, and eight main sessions) which each lasted about 70 minutes.

The following analyses focus on the eight main sessions. In total, 192 subjects took part in these sessions. Apart from one subject who tried to use a cell phone calculator to solve the task, all participants with a performance of more than 20 correctly solved items per round were excluded from the analysis to eliminate outliers. This leaves 189 subjects (96 men and 93 women) for the analysis.

First of all, I check for pre-treatment differences between the treatment groups by performing Kruskal-Wallis tests on observable characteristics to test whether the samples are drawn from the same distribution. The tests detect no significant differences on observables except for biased beliefs about others' performance in rounds one and two (χ^2=8.24, p=.0413 and χ^2=11.56, p=.0009).[13] However, this means that results with respect to beliefs about others' performance can only be interpreted with caution.

In addition, I assess whether there is productivity sorting of the reference group (i.e.,

[12] Again, subjects were given detailed instructions and had to answer several control questions to make sure that they understood that they would maximize their expected payoff by choosing their subjective probability of winning as the threshold value.

[13] The averages of observable characteristics (age, risk aversion, performance and beliefs in rounds one and two) can be found in Table 5 in Appendix B.

whether the performance of the selected sample is indeed better than the performance of the random sample) by calculating a Kolmogorov-Smirnov test for equality of distribution functions between the groups, which turns out to be significant (D=0.35, p=.086), indicating that there is indeed productivity sorting in my sample. Individuals facing the selected sample should factor that in when making competition decisions.[14]

Having established these preliminaries, I now turn to investigating the effects of overconfidence and competition neglect in my sample. I do so in two steps: First, I focus on beliefs, give an impression of how they look like in my sample and show how they are affected by the treatment variations. Then, I turn to entry decisions by again first presenting the general pattern and then addressing treatment effects.

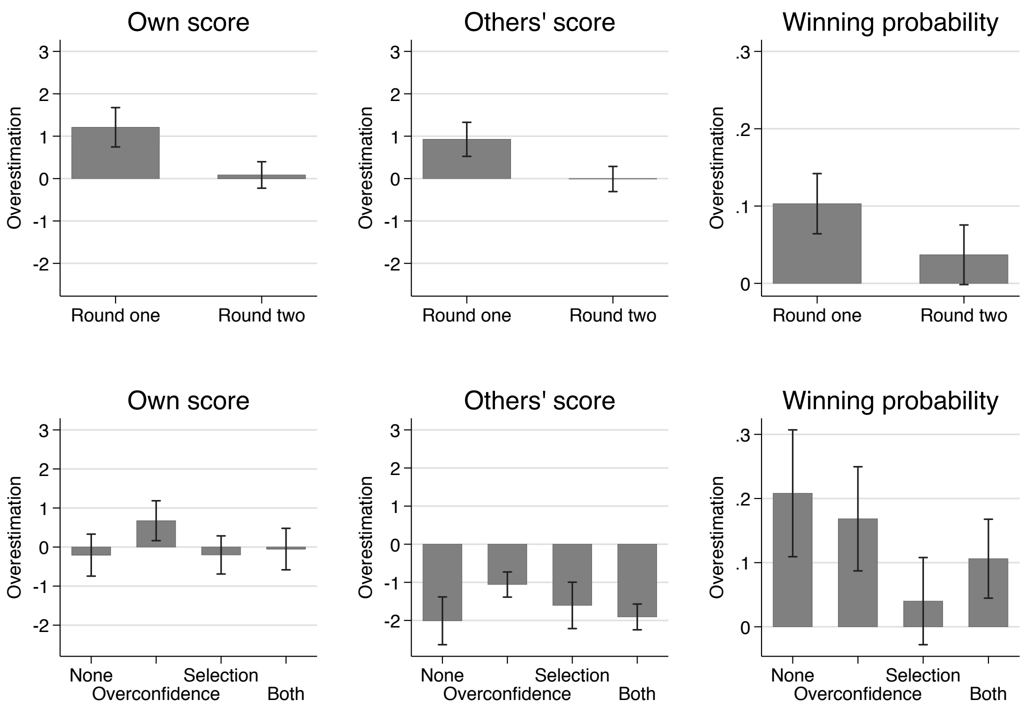
5.1. Beliefs

In the following paragraphs, I take a closer look at participants' beliefs across rounds. Despite having elicited whole belief distributions, I first calculate expected values of performances and focus on these. To arrive at a measure for subjects' overconfidence, I then compute the difference between expected and actual number of correctly solved items. Positive values indicate that own performance is overestimated, while negative values indicate that it is underestimated. Analogously, the difference between expected and actual number of correctly solved items of a randomly drawn individual from the competition group is calculated. For subjective winning probability, I compute the actual probability of winning for every number of correctly solved items based on reference group performance and subtract this from participants' stated winning probability. This procedure results in three values for erroneous beliefs per round.

First, I explore how accurate beliefs are and how they are distributed in my sample. To this end, Figure 2 depicts average deviations in beliefs from actual scores. Here, two things are especially noteworthy. First, the upper row clearly shows that on average, participants in my sample have very accurate beliefs by round two: while there is overestimation of own and (less drastically) also of others' performance as well as

[14] Note that productivity sorting can also be seen in the main sessions (D=0.38, p<.001).

Figure 2: *Average bias in beliefs*

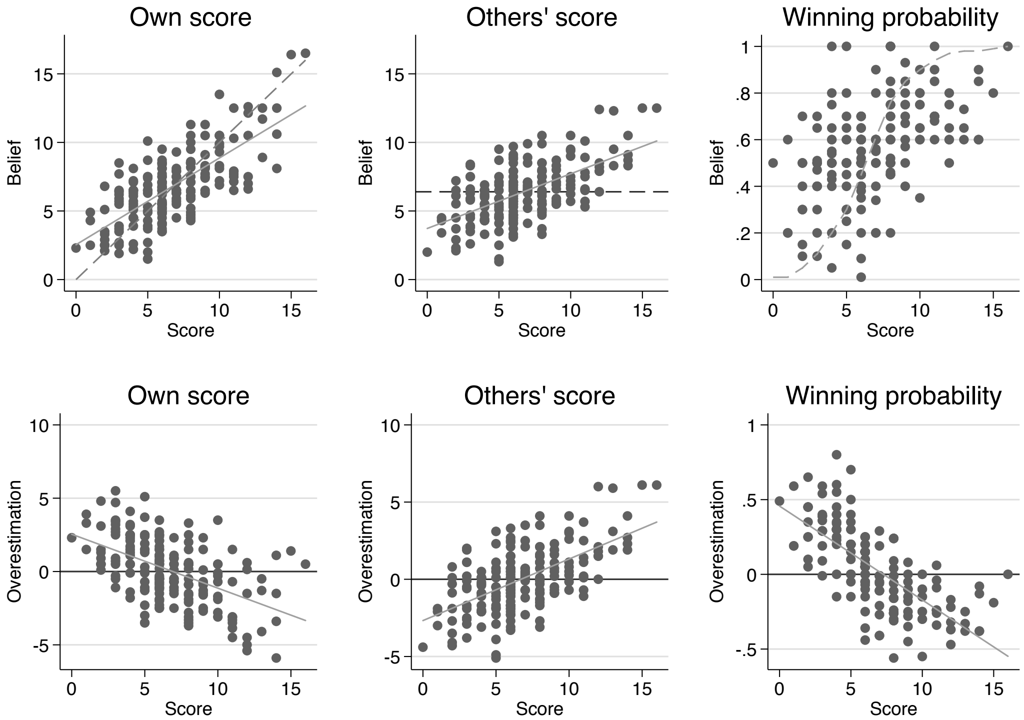


Notes: Upper row depicts average bias in beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) in rounds one and two in the main sessions. Lower row depicts average bias in beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) by Treatment in round three in the main sessions. Calculations are based on 189 observations. Means are depicted as bars, 95% confidence intervals as error bars.

overestimation of the chances of being better than a randomly drawn individual from the reference group in the first round, none of these statements holds true for average beliefs in round two. In addition, the lower row of Figure 2 shows that in round three, there are differences in beliefs between treatments: underestimation of others' performance and overestimation of winning probability is stronger when facing the selected than when facing the random competition group, providing evidence for competition neglect. Interestingly, this difference is more pronounced when no feedback is provided.

Next, I explore heterogeneity in beliefs. As can be seen in Figure 3, there is systematic heterogeneity even in round two. Plotting beliefs and overestimation values by actual performance, it is evident that low-performing subjects are more prone to overestimate themselves while underestimating the reference group as well as overestimating their

Figure 3: *Beliefs and overestimations by actual performance in round two.*



Notes: Upper row of figure depicts beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) depending on actual performance in round two of the main sessions. Lower row depicts bias in beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) depending on actual performance. Calculations are based on 189 observations.

chances of performing better than a randomly drawn individual from the reference group.[15]

Having established how participants' beliefs behave, I now investigate whether they are influenced by giving feedback and by the composition of the reference group to see whether the treatments worked as expected. This way, exploring changes in subjects' assessments serves as a prerequisite for identifying the effect of competition neglect and

[15] One might argue that this effect is at least partly mechanic as high-performing individuals have little scope to overestimate and much scope to underestimate themselves, while for low-performing individuals, the opposite is true. However, this does not apply to beliefs about others as these should be independent of own performance. In addition, the relationship between actual and overestimation of own performance and winning probability is still significantly negative if overestimation is not considered continuously, but as binary variable (point-biserial correlations are r=–.37, p<.001 for own performance, and r=–.66, p<.001 for winning probability).

overconfidence on choices. Thus, I explore determinants of changes in beliefs between rounds two and three in various specifications in Table 2.

First, to see whether competition neglect is reflected in participants' beliefs, focus on the coefficient for the selected competition group when change in winning probability is the dependent variable, which is insignificant regardless of specification; thus, compound beliefs are not adjusted more strongly when facing the selected competition group.

Second, I investigate whether absolute and relative overconfidence are ameliorated by providing feedback. The first thing to note is that the coefficient for feedback is insignificant regardless of specification. However, since it might be important to account for the fact that subjects receiving feedback have better information, I additionally look at the interaction between feedback and round two performance. As can be seen in columns (4) to (6) of Table 2, the positive effect of previous performance on belief changes is indeed higher when performance feedback is provided. In columns (7) to (9) of Table 2, I include round two overestimations and interactions with feedback to explore whether subjects exhibiting higher pre-treatment overestimation correct their estimations more strongly. This would mean that changes in beliefs are sensible, that is, feedback ameliorates perception biases. Indeed, round two overestimations of own and others' performance have a negative influence on changes in beliefs, meaning that stronger round two overestimations are associated with changes towards lower performance beliefs. When feedback is provided, this negative influence is even stronger. For winning probability, previous overestimation only has a significantly negative influence on belief changes when feedback is provided.

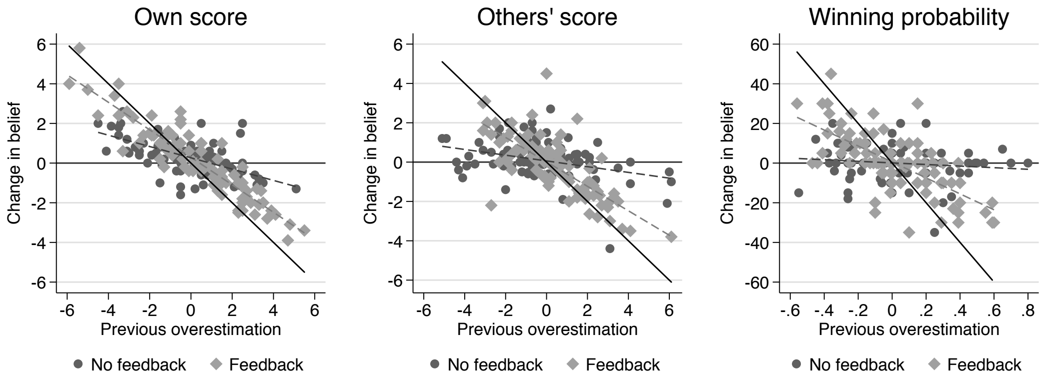
The benefiting role of performance feedback on performance beliefs is illustrated in Figure 4. Here, it can be seen clearly that while there is some adjustment in beliefs, it is not perfect (if it was, observations would lie on the inverse 45 degree line). For overestimation of own and others' performance, there is adjustment even without feedback (dark dashed lines do not overlap with horizontal line), while this is not true for winning probability. More importantly, the adjustment is stronger if feedback is provided, as the comparisons of dark and light dashed lines in all three graphs indicate. Thus, providing feedback leads to sensible changes in beliefs.

Table 2: *Determinants of changes in beliefs between rounds two and three*

DV: Changes in beliefs about...	(1) Self	(2) Others	(3) Probability	(4) Self	(5) Others	(6) Probability	(7) Self	(8) Others	(9) Probability
Feedback = 1	-0.237	-0.225	-0.359	-0.292	-0.195	-1.013	0.0129	0.00671	0.446
	(0.200)	(0.193)	(1.698)	(0.191)	(0.188)	(1.628)	(0.115)	(0.147)	(1.413)
Selection = 1	-0.140	-0.0784	-1.592	-0.196	-0.0485	-2.260	0.0113	0.200	-3.088**
	(0.193)	(0.190)	(1.699)	(0.186)	(0.183)	(1.646)	(0.107)	(0.133)	(1.385)
Previous Performance	0.538***	-0.160	3.230***	0.257**	-0.0101	-0.106			
	(0.119)	(0.112)	(0.941)	(0.103)	(0.114)	(0.795)			
1.feedback				0.590**	-0.315	7.018***			
#c.Score_pre				(0.233)	(0.225)	(1.473)			
Previous overestimation							-0.294***	-0.143***	-3.118
							(0.0382)	(0.0504)	(3.552)
1.feedback							-0.401***	-0.511***	-37.89***
#c.overestimation_pre							(0.0529)	(0.0719)	(5.375)
Observations	187	187	187	187	187	187	187	187	187
Adjusted R^2	0.136	0.00727	0.0783	0.175	0.0172	0.159	0.730	0.462	0.365

Notes: Table reports results of OLS regression. Controls include risk aversion, gender, age, an indicator for field of study, and math grade. Previous score is standardized. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure 4: *Changes in beliefs by overestimation in round two*



Notes: Figure depicts changes in beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) between rounds two and three depending on overestimation in round two of the main sessions. Calculations are based on 189 observations.

To sum up, the analysis of individuals' beliefs demonstrates that there is systematic heterogeneity in beliefs about own performance, others' performance, and winning probability. While changes in beliefs are not influenced by the composition of the reference group, revealing competition neglect, they are positively affected by performance feedback, revealing overconfidence that can be ameliorated. Hence, these results establish that both of the mechanisms that are discussed to influence competition entry are prevalent in this experiment.[16]

5.2. Competition Decisions

After having explored individuals' beliefs, I now turn to the main dependent variable, decisions to compete, and thus to the main focus of this paper, namely how overconfidence and competition neglect influence selection into competition. Again, I investigate this question by first providing a general impression of competition decisions in my sample before systematically investigating determinants of competition choices.

To start, I assess whether there is excess entry in my setup. Across all treatments, 33.33% of individuals choose the competitive incentive scheme. Considering the actual

[16] Note that in contrast to what other studies find, there are no gender differences in beliefs. For the interested reader, analyses of gender differences are provided in Appendix C.

distribution of performances in the reference group, a payoff-maximizing individual should choose competition if her round three performance equals at least eight correctly solved items.[17] Assuming that performances in rounds two and three are equal, 35.98% of individuals have higher expected earnings when choosing competition. Thus, I do not find excess entry compared this benchmark in my sample.

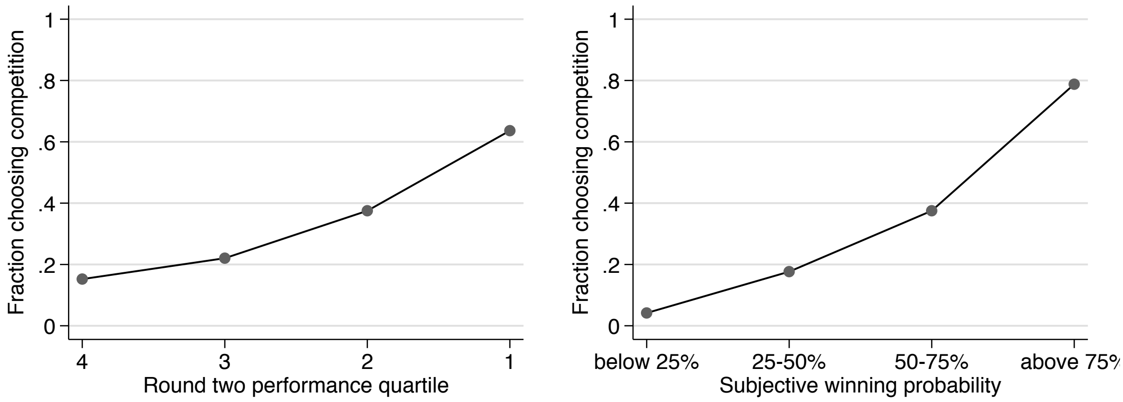
However, these comparisons are based on the assumption of payoff-maximization and thus are rather strict. Therefore, I next explore competition decisions depending on round two performance and beliefs about winning probability. The left panel of Figure 5 depicts the proportion of participants selecting into competition by performance quartile in the initial competition. Contrary to what Niederle and Vesterlund (2007) find, there is a positive relationship between performance quartile and proportion of competition decisions. The right panel of Figure 5 depicts the proportion of participants selecting into competition by subjective probability of winning. There is a positive relationship between beliefs and the proportion of subjects selecting into competition that is even stronger than for previous performance. Thus, individuals' competition decisions in my sample are closely tied to their previous performance and beliefs.

To answer my central question whether competition decisions are influenced by overconfidence and competition neglect, I explore competition decisions by treatment; if competition neglect plays a role, differences between the *None* and *Selection* as well as between the *Overconfidence* and *Both* treatments should be smaller than expected. If errors in performance beliefs influence decisions, this should be reflected in the difference between treatments with and without feedback as well as in the influence of beliefs. The fraction of individuals selecting the competitive incentive scheme by treatment as well as the fraction of individuals for which it is the payoff-maximizing choice are depicted in Figure 6. Apart from the fact that absolute ratios of subjects deciding to compete seem pretty low when compared to previous literature[18] and are lower

[17] Figure 8 in Appendix B plots actual chances of winning depending on performance. It can be seen that the winning probability exceeds 50% for more than seven correctly solved items.

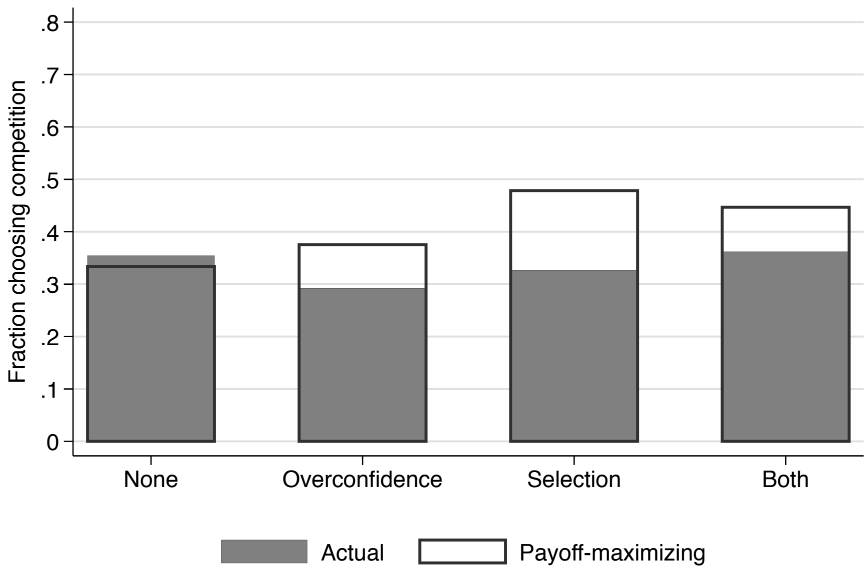
[18] In Niederle and Vesterlund (2007), 73% of men and 35% of women enter. While their setup is slightly different from mine as participants have to outperform three others, studies also employing the two-person competition with a randomly drawn competitor find similar entry rates; for instance, in Berlin and Dargnies (2016), 63.2% of men and 35.1% of women initially chose to compete.

Figure 5: *Competition choices by previous performance and beliefs*



Notes: Figure depicts fractions of individuals choosing the competitive incentive scheme in round three of the main sessions by performance quartile in round two (left panel) and subjective winning probability (right panel). Calculations are based on 189 observations.

Figure 6: *Competition choices by treatment*



Notes: Figure depicts fraction of participants choosing the competitive incentive scheme in round three of the main sessions by treatment condition. Calculations are based on 189 observations.

than the payoff-maximizing benchmark, there are no striking differences in competition rates between the treatments. Testing for gender differences within treatment and for treatment differences with Fisher exact tests does not detect significant effects.

Next, I take a closer look at determinants of competition decisions. The results of linear probability model estimations with decisions to compete as dependent variable are shown in Table 3. First, it can be seen that the main effects for performance feedback as well as for facing the self-selected competition group are nonsignificant, which is not surprising given Figure 6. Second, as indicated by Figure 5, previous performance has a positive effect on the probability to choose the competitive incentive scheme. Additionally, while the gender and competition literature has established that there is a gender difference in competitiveness (Niederle and Vesterlund, 2011), I find no effect of gender in my sample, while risk aversion has a significantly negative effect on selection into competition.[19] In columns (3) and (4), performance beliefs are included in the model. Now, the coefficients for all three performance assessments are significant, indicating that these beliefs influence competition decisions beyond performance. More specifically, holding actual round two performance constant, participants are more likely to select the competitive incentive scheme if they believe to have a higher performance, if they assess their competitors' performance to be lower, and if they believe to have a higher probability of winning. Thus, overly optimistic beliefs indeed influence participants' competition decisions.

Regarding the central question asked in this paper, the following conclusions can be drawn based on the experimental results: as the proportion of participants selecting into the competitive incentive scheme does not depend on the composition of the competition group, competition neglect increases entry into competition. In addition, I also find evidence for the positive influence of absolute and relative overconfidence on entry into competition. Performance feedback does ameliorate bias in performance beliefs, but does not influence competition decisions directly.

[19] A separate analysis of treatment effects for men and women is provided in Appendix C.

Table 3: *Determinants of competition choices*

DV: Competition (Y/N)	(1)	(2)	(3)	(4)
Feedback = 1	-0.00974	0.00317	0.00333	0.00785
	(0.0634)	(0.0656)	(0.0591)	(0.0598)
Selection = 1	-0.0519	-0.0582	-0.0311	-0.0357
	(0.0627)	(0.0665)	(0.0567)	(0.0584)
Previous performance	0.0581***	0.0613***	-0.0266	-0.0255
	(0.0108)	(0.0114)	(0.0235)	(0.0234)
Own performance belief			0.107***	0.112***
			(0.0306)	(0.0313)
Others' performance belief			-0.0797***	-0.0799***
			(0.0271)	(0.0291)
Subj. winning probability			0.628***	0.620***
			(0.168)	(0.176)
Female = 1	-0.0346	-0.00442	0.0294	0.0538
	(0.0659)	(0.0694)	(0.0587)	(0.0606)
Risk Aversion	-0.0438**	-0.0424**	-0.0305*	-0.0291*
	(0.0176)	(0.0175)	(0.0163)	(0.0166)
Constant	0.259*	0.750***	0.112	0.433
	(0.156)	(0.262)	(0.190)	(0.276)
Controls	No	Yes	No	Yes
Observations	189	187	189	187
Adjusted R^2	0.153	0.151	0.306	0.304

Notes: Table reports results of linear probability estimations. Robust standard errors in parentheses. Estimations are based on round three. Controls include age, an indicator for field of study (business, economics, STEM, humanities, law, psychology, other social sciences, other), and math grade. *** p<0.01, ** p<0.05, * p<0.1.

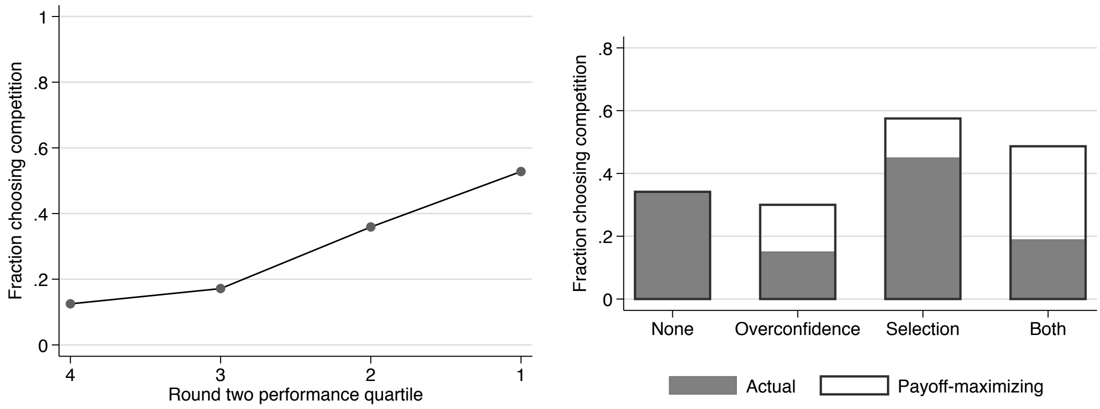
5.3. The Effect of Pre-Treatment Belief Elicitation

In light of the vast literature on competitiveness, where competition rates are excessive and overconfidence is prevalent on average, it is surprising that in my setup, belief calibration works reasonably well, and that competition decisions are tied closely to performance and expectations. However, one clear point of departure of the present study is that the belief elicitation is conducted with three questions and that the belief questions are asked before the task in every round, making them salient. In this way, participants are forced to think very carefully about how they perceive their performance in the task relative to others, while other studies rely on eliciting the beliefs only after participants have made their decisions. Put differently, the belief elicitation procedure itself could cause subjects to think deliberately about the aspects relevant to the decision while not asking allows them to make intuitive "gut decisions". This is in line with dual-process theories of higher cognition in psychological research (see Kahneman, 2011; Evans, 2008; Evans and Stanovich, 2013).

To address the fact that competition decisions in the main sessions looked starkly different from the findings in the literature and to explore whether the belief elicitation procedure in itself might already change the way the competition decision is approached, I conducted eight additional sessions, yielding a total of 158 observations (88 women and 70 men). These sessions have the exact same setup as the main sessions, with the exception that the belief elicitation in rounds one and two is dropped so that beliefs are now elicited for the first time only after the competition choice is made.

Of course, this change in procedure makes it impossible to explicitly investigate changes in beliefs as the pre-decision measures are missing. However, I am still able to investigate the relationship between competition decisions and performance and to compare competition decisions across treatments. These relationships are illustrated in Figure 7. As can be seen in the left panel, exploring competition choices depending on previous performance reveals a slightly different picture than in the main sessions: there still is a positive relationship between performance quartile and the proportion of individuals selecting into competition, but this relationship seems to be less pronounced.

Figure 7: *Competition choices in additional sessions*



Notes: Figure depicts fraction of men and women choosing the competitive incentive scheme in round three of the additional sessions by round two performance quartile (left panel) and by Treatment (right panel). Calculations are based on 25 men and 44 women.

In a next step, the fraction of individuals selecting into competition depending on the treatment conditions is depicted in the right panel of Figure 7. Again, there is less entry than would be payoff-maximizing, and the general pattern suggests that overconfidence as well as competition neglect influence choices. While there is a slight increase in the fraction of individuals selecting into competition both from *None* to *Selection* and from *Overconfidence* to *Both*, this increase is very small (and much smaller than the optimal change measured by payoff-maximizing choices would be). Most interestingly, in contrast to the main sessions, there is a sharp drop in the fraction of competition choices when feedback is provided. Testing for treatment differences in this sample reveals that, unlike in the main sessions, the fraction of individuals selecting into competition differs between treatments (two-sided Fisher exact test: p=.011).

Next, I again take a closer look at determinants of competition decisions by estimating linear probability models. The results are shown in Table 4. In contrast to the estimations based on the sessions with belief elicitation in each period, providing feedback has a significant, negative effect on the probability to choose the competitive incentive scheme in models (1) and (2). Abstract from this difference, the remainder of the results from the estimations is similar to the main sessions: The main effect for selection is again nonsignificant. Previous performance has a positive effect on competition

Table 4: *Determinants of competition choices (additional sessions)*

DV: Competition (Y/N)	(1)	(2)	(3)	(4)
Feedback = 1	-0.197***	-0.215***	-0.111	-0.126*
	(0.0656)	(0.0694)	(0.0708)	(0.0755)
Selection = 1	-0.0516	-0.0472	-0.0373	-0.0411
	(0.0656)	(0.0700)	(0.0628)	(0.0656)
Previous performance	0.0394***	0.0333**	0.00289	-0.00125
	(0.0126)	(0.0139)	(0.0166)	(0.0174)
Own performance belief			0.0665***	0.0654***
			(0.0220)	(0.0225)
Others' performance belief			-0.0668***	-0.0699***
			(0.0227)	(0.0237)
Subj. winning probability			0.387*	0.456*
			(0.211)	(0.241)
Female = 1	-0.110	-0.0990	-0.0543	-0.0450
	(0.0683)	(0.0738)	(0.0665)	(0.0702)
Risk Aversion	-0.0386**	-0.0372*	-0.0361*	-0.0382*
	(0.0193)	(0.0208)	(0.0192)	(0.0202)
Constant	0.442***	0.378*	0.387*	0.321
	(0.150)	(0.207)	(0.197)	(0.273)
Controls	No	Yes	No	Yes
Observations	158	157	158	157
Adjusted R^2	0.156	0.122	0.238	0.212

Notes: Table reports results of linear probability estimations. Robust standard errors in parentheses. Estimations are based on round three. Controls include age, an indicator for field of study (business, economics, STEM, humanities, law, psychology, other social sciences, other), and math grade. *** p<0.01, ** p<0.05, * p<0.1.

propensity, but only significantly so until performance beliefs are included in columns (3) and (4). Risk aversion has a significantly negative effect on competitiveness, while there is again no main effect for gender.

Finally, to investigate whether the results obtained in these additional sessions are indeed closer to the competition patterns that are found in the literature, I compare selection into competition by men and women. Overall, 37.14% of men choose the competitive incentive scheme, while only 21.59% of women do, making the difference in choices more extreme compared to the sessions with belief elicitation in every round. Indeed, the difference in competition choices between men and women is significant for the additional sessions (two-sided Fisher exact test: p=.035). In *Selection*, which is closest to the setting used in the gender and competition literature, the difference is even more pronounced: 66.67% of men choose to compete while only 27.27% of women do.[20] Indeed, when testing for gender differences separately for each of the treatments, it becomes evident that the overall gender difference in choices is driven by *Selection* (two-sided Fisher exact test: p=.024), as the difference is insignificant in the other three treatments.

To sum up, the additional sessions reveal that without belief elicitation before the competition choice, feedback has a significant influence on choices and that gender differences are relatively more pronounced. These results show that the subject pool and setup used is comparable to other studies in the literature, thus enhancing confidence in the validity of the main analysis. In addition, the results from the additional session reveal that the belief elicitation procedure itself might lead to changes in the way individuals process the decision, thus highlighting potentially interesting additional directions for research of how to "debias" people in these choice settings.

[20] This is quite close to the 73% and 35% in Niederle and Vesterlund (2007) and the 63.2% and 35.1% in Berlin and Dargnies (2016).

6. Discussion

In this paper, I presented the results of a laboratory experiment designed to explore two distinct mechanisms underlying perception biases that influence selection into competition: overconfidence with respect to own and others' performance and competition neglect. In my setting, I find that there is systematic heterogeneity in perception biases, with low-performing individuals overestimating their own performance and their chances of success while underestimating performance in the competition group, while the opposite is true for high-performing individuals. These biases in performance beliefs are ameliorated by feedback; however, individuals persistently disregard the composition of the competition group they face. Investigating determinants of entry decisions to tackle the key question of this paper whether competition neglect and overconfidence influence entry into competition, I find that both matter. However, decisions are closely tied to previous performance and assessments, and there are no gender differences.

Based on these findings, I argue that eliciting subjects' beliefs pre-treatment might have changed their mode of thinking, inducing them to think about the relevant evaluations more carefully rather than making intuitive decisions. Results from additional sessions where I refrain from eliciting subjects' beliefs pre-treatment are closer to the existing literature: in these sessions, there is a much more pronounced difference between competition choices of men and women, and choices are not as closely tied to performance.

However, this comparison implies that in the main sessions, competition neglect persists even though the belief elicitation procedure points individuals to the fact that they should evaluate their own as well as competitors' performance and factor those into their chances of winning. While this changes their choices and ties them closer to performance, it does not help to address competition neglect, which appears to be a persistent phenomenon. Hence, educating individuals about the aspects that matter for their decision does lead to less excessive entry, but does not ameliorate competition neglect. To address this bias, other, more direct ways of educating decision makers have to be applied. While this paper has mainly established competition neglect as a relevant bias in a laboratory setting, it is up to future research to come up with ways to specifically address this bias and to explore its impact in real-world settings.

On a more general note, based on these results, it seems important to keep in mind the mode of thought in which competition decisions are made. One might argue that while decisions made in the laboratory involve small stakes and are thus more prone to be gut decisions, decisions like career choices naturally involve more careful thought and are thus more closely tied to actual performance. Hence, laboratory findings might overestimate the differential in competitiveness between women and men. However, to be able to make detailed statements about the importance of the role of the mental resources deployed for decision making, one would have to design a mechanism experiment that explicitly manipulates the mode of processing. I leave this for future research.

Appendix

A. Proofs

Existence of q^*

Proof. First, it is straightforward to see that the denominator is always > 0. In addition, keeping in mind that $\xi > 1$ and $q_j \in [0;1]$, one can see that the numerator is ≥ 0 as long as $s \leq 2$ (and thus, $q^* \geq 0$ as well): $2-(1-q_j)s \geq 0 \Leftrightarrow s \leq \frac{2}{1-q_j}$. This condition is strictest if $q_j = 0$, when it becomes $s \leq 2$. On the other hand, the numerator is smaller or equal than the denominator (an, in turn, $q^* \leq 1$) as long as $s \geq \frac{2}{2-q_j}$. This is strictest for $q_j = 1$, when it becomes $s \geq 2$. Thus, for $s = 2$, $q^* \in [0;1]$ for all possible values of q_j (and ξ). $\square$

Uniqueness of q^*

Proof. First, note that due to the common knowledge assumption, q^* is symmetric. For uniqueness, that is, for q^* to be a global threshold, one needs to show that $\forall i \in N$, $q_i' < q^* \Leftrightarrow \mathbf{E}_{NC}(q_i') > \mathbf{E}_C(q_i')$ and $q_i'' > q^* \Leftrightarrow \mathbf{E}_{NC}(q_i'') < \mathbf{E}_C(q_i'')$. To see this, plug q_i' into $\mathbf{E}_{NC} > \mathbf{E}_C$. Solving for q_i' yields $q_i' < \frac{2-(1-q_j)s}{\xi[s(2-q_j)-2]-[(1-q_j)s-2]}$, which equals $q_i' < q^*$. Similarly, plugging q_i'' into $\mathbf{E}_{NC} < \mathbf{E}_C$ and solving for q_i'' yields $q_i'' > \frac{2-(1-q_j)s}{\xi[s(2-q_j)-2]-[(1-q_j)s-2]}$, which equals $q_i'' > q^*$. Thus, q^* is unique. $\square$

Comparative statics of q^*

$$\frac{\partial q^*}{\partial s} = -\left(\frac{(1-q_j)}{\xi[s(2-q_j)-2]-[(1-q_j)s-2]} + \frac{[2-(1-q_j)s] \times (\xi(2-q_j)-(1-q))}{\{\xi[s(2-q_j)-2]-[(1-q_j)s-2]\}^2}\right)$$

To see that this is < 0, note that the denominators of both fractions are > 0. Then, the first fraction is ≥ 0 because $(1-q_j) \geq 0$. The numerator of the second fraction consists of two parts. As established in the existence proof, the first, $[2-(1-q_j)s]$, is positive. The second, $(\xi(2-q_j)-(1-q))$, is > 0 as long as $\xi > \frac{1-q}{2-q}$, which is satisfied for $\xi > 0.5$. Thus, the whole second fraction is > 0. Finally, the whole expression is < 0 (and not ≤ 0) because even if the first fraction $= 0$, which happens for $q_j = 1$, the second fraction is > 0.

$$\frac{\partial q^*}{\partial \xi} = -\frac{[2-(1-q_j)s] \times [s(2-q_j)-2]}{\{\xi[s(2-q_j)-2]-[(1-q_j)s-2]\}^2}$$

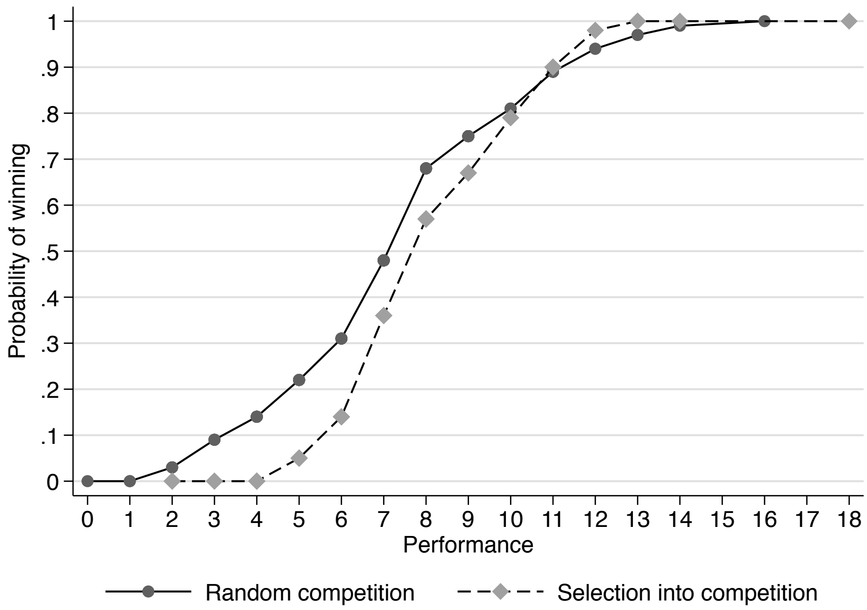
To see that this is < 0, note that the denominator is > 0. The first part of the numerator is again > 0, as established in the existence proof. The second part of the numerator is positive as long as $s \geq \frac{2}{2-q_j}$, making the whole expression negative.

$$\frac{\partial q^*}{\partial q_j} = \frac{s}{\xi[s(2-q_j)-2]-[(1-q_j)s-2]} + \frac{[2-(1-q_j)s] \times (\xi-1)s}{\{\xi[s(2-q_j)-2]-[(1-q_j)s-2]\}^2}$$

Again, both denominators are > 0. Thus, the first fraction is also > 0 as s is positive by definition. The numerator of the second fraction again consists of two parts, with $[2-(1-q_j)s]$ being positive. As $\xi > 1$, the second part is positive as well, which means that the whole expression is > 0.

B. Figures and Tables

Figure 8: *Actual chance of winning by round three performance*



Notes: Figure depicts the actual chances of winning given reference group performance when the competition group consists of the whole reference group (solid line) and when the competition group consists of self-selected individuals only (dashed line). Calculations based on 48 individuals in the reference group, of which 21 chose the competitive incentives.

Table 5: Balancing table

	None	Overcon-fidence	Selection	Both	Chi^2
Gender	0.50	0.48	0.50	0.49	0.23
	(0.51)	(0.50)	(0.51)	(0.51)	
Age	23.69	23.17	23.30	24.15	3.03
	(3.07)	(4.70)	(3.60)	(5.75)	
Risk aversion	5.67	5.94	6.30	6.38	3.34
	(2.00)	(2.23)	(1.76)	(1.55)	
Round 1					
Score	5.73	6.27	6.30	5.55	3.62
	(2.95)	(2.81)	(2.74)	(2.88)	
Overestimation of self	1.23	1.79	0.32	1.47	5.24
	(2.96)	(3.61)	(2.70)	(3.49)	
Overestimation of others	0.66	1.88	0.50	0.64	8.24**
	(3.03)	(2.97)	(2.42)	(2.56)	
Overestimation of probability	0.14	0.04	0.07	0.16	5.76
	(0.29)	(0.26)	(0.28)	(0.24)	
Round 2					
Score	6.40	7.08	6.93	6.32	2.55
	(3.06)	(3.00)	(3.06)	(3.14)	
Overestimation of self	-0.16	0.51	-0.15	0.13	3.69
	(2.17)	(2.28)	(1.72)	(2.44)	
Overestimation of others	-0.49	0.75	-0.13	-0.17	11.56***
	(2.33)	(1.85)	(2.12)	(1.77)	
Overestimation of probability	0.08	-0.01	0.01	0.07	3.29
	(0.30)	(0.25)	(0.25)	(0.27)	
Observations	48	48	46	47	

Notes: Table reports variable means in the main sessions by treatment. Standard deviations in parentheses. Estimations are based on round three. Column "Chi^2" reports Kruskal Wallis test statistics / Chi^2 test statistic for binary variable. *** p<0.01, ** p<0.05, * p<0.1.

C. Gender Differences

In this section, the analyses of beliefs and competition decisions are done separately for men and women to investigate whether the established difference from the literature are prevalent in my sample as well and whether the influence of competition neglect and overconfidence on competition decision is different for men and women.

Beliefs

Figures 9 and 10 are similar to Figures 2 and 3 in Section 5 of Chapter 4, but depict beliefs split up by men and women. It can be seen that the patterns for men and women are more or less the same, and that there are are no significant differences between men and women in overconfidence in my sample. This is surprising as previous literature on gender differences in competitive environments has established heterogeneity in overconfidence as one important driver of gender differences in competitiveness. For instance, Niederle and Vesterlund (2007) find that 75% of men and 43% of women believe to have the highest performance in a four-person group.

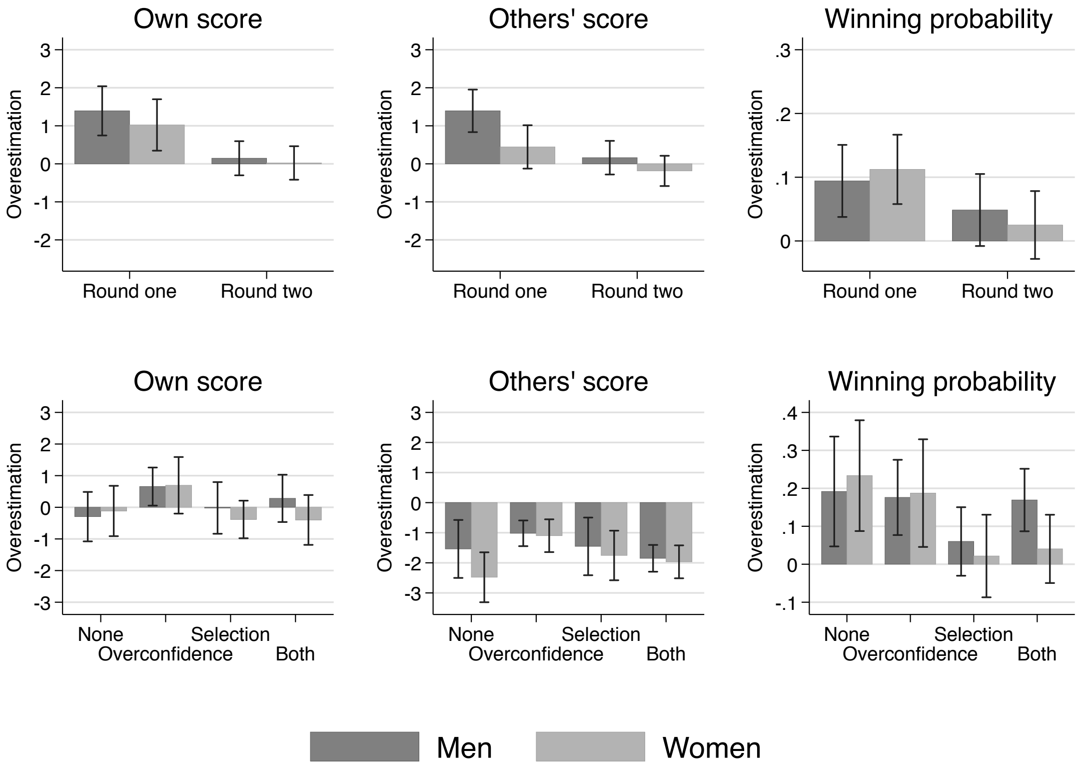
Competition Decisions

Now, I analyze how competition decisions and the effects of competition neglect and overconfidence differ for men and women.

An important condition to being able to compare competition decisions of men and women directly is that there are no gender differences in performance – in this case, the money-maximizing choice is the same for both genders and, normatively speaking, competition decisions should not differ. This is also an important assumption underlying arguments for acting against gender inequality in labor markets and calls for an increased number of women in competitive professions. Thus, Figure 11 depicts average performance by round and gender in the main sessions of the experiment. While in most papers, no significant difference in performance between men and women is found for this task, I find that in my sample, men on average solve 1.28 items more than women in round one, 0.98 items more in round two, and 0.80 items more in round three. Thus, they perform significantly better than women in all rounds of the experiment (two-sided Mann-Whitney tests yield z=2.43, p=.0148, z=1.98, p=.0476, and z=1.75, p=0.0803).[21]

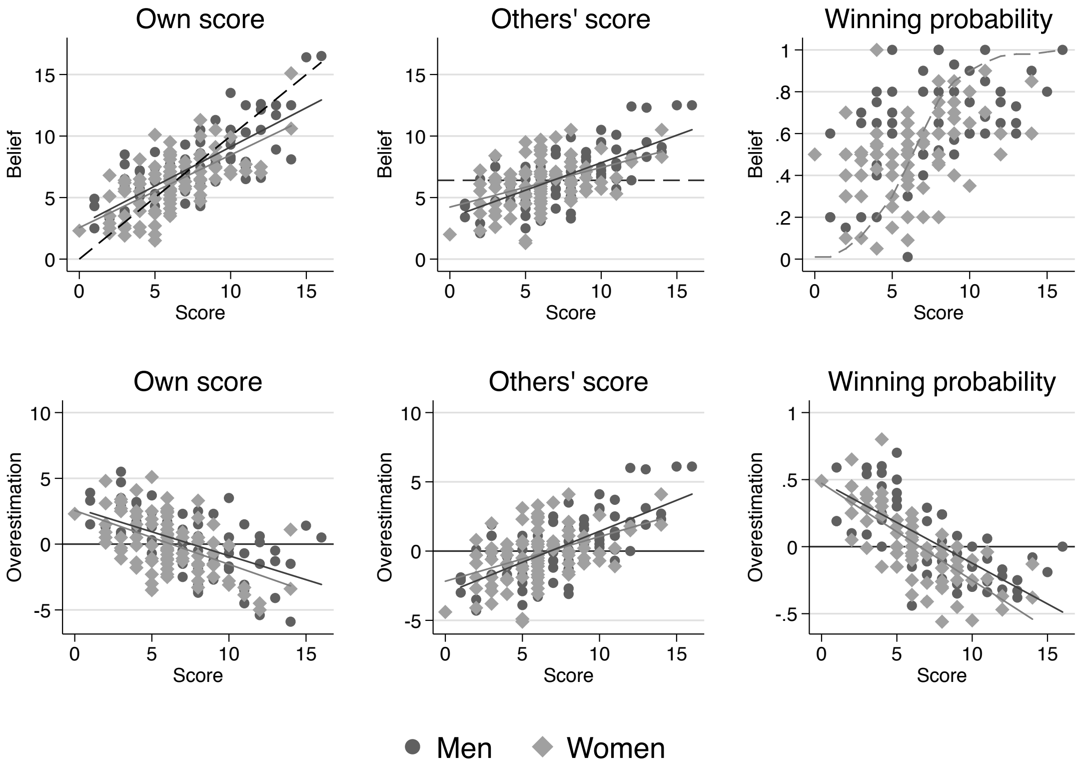
[21] Note that while some studies have indicated that there might be an adverse effect of having to

Figure 9: *Average bias in beliefs*



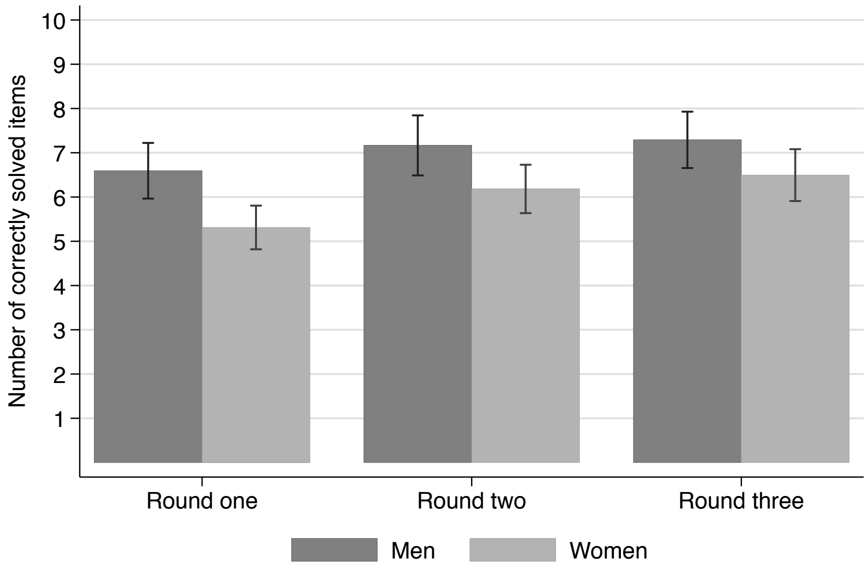
Notes: Upper row depicts average bias in beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) in rounds one and two in the main sessions. Lower row depicts average bias in beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) by Treatment in round three in the main sessions. Calculations are based on 189 observations. Means are depicted as bars, 95% confidence intervals as error bars.

Figure 10: *Beliefs and overestimations by actual performance in round two.*



Notes: Upper row of figure depicts beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) depending on actual performance by men and women in round two of the main sessions. Lower row depicts bias in beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) depending on actual performance by men and women in round two of the main sessions. Calculations are based on 96 men and 93 women.

Figure 11: *Average performance by gender across rounds*



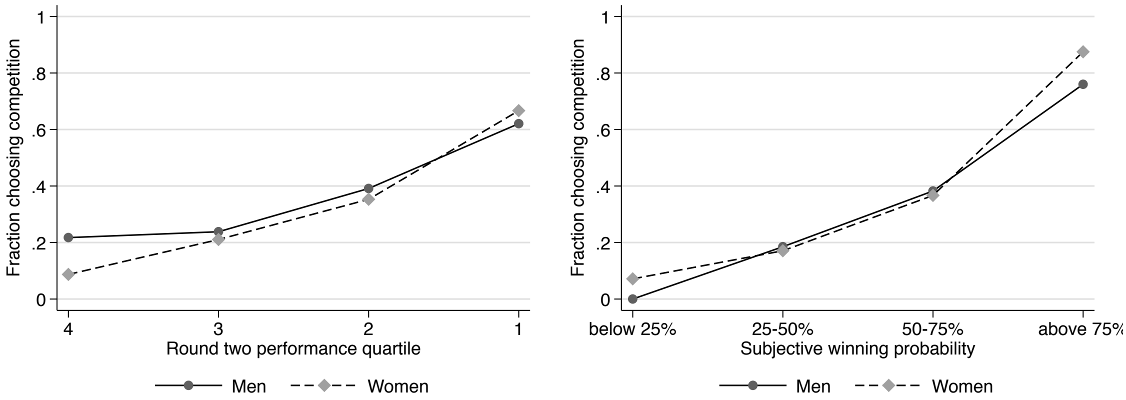
Notes: Figure depicts average number of correctly solved items by men and women across rounds in the main sessions. Calculations are based on 96 men and 93 women. Means are depicted as bars with 95% confidence intervals.

Next, consider entry rates. When considering average entry across treatments, contrary to Niederle and Vesterlund (2007) and most of the literature, there exists no significant gender gap: 38.54% of men and 27.96% of women chose the competitive incentive scheme (two-sided Fisher's exact test yields p=.165).

Figure 12 depicts the proportion of participants selecting into competition by performance quartile in the initial competition. Contrary to what Niederle and Vesterlund (2007) find, there is a positive relationship between performance quartile and proportion of individuals selecting into competition for both men and women. In addition, the relationship looks rather similar for men and women, with the exception of subjects in the lowest quartile: among these, competition rates for men are higher than for women, indicating that there is more overcompetition of low-performing men than women. The right panel of Figure 12 depicts the proportion of men and women selecting into competition by their subjective probability of winning. There is a positive relationship between beliefs and the proportion of subjects selecting into competition

work under competitive incentives for women Gneezy et al. (2003), I find the opposite. Of course I cannot disentangle these effects from learning; however, the fact that womens' average performance improves more strongly between rounds one and two than mens' average performance speaks against a strong negative effect and a strong positive of the competitive incentives for women and men, respectively.

Figure 12: *Competition choices by previous performance and beliefs*



Notes: Figure depicts fractions of men and women choosing the competitive incentive scheme in round three of the main sessions by performance quartile in round two (left panel) and subjective winning probability (right panel). Calculations are based on 96 men and 93 women.

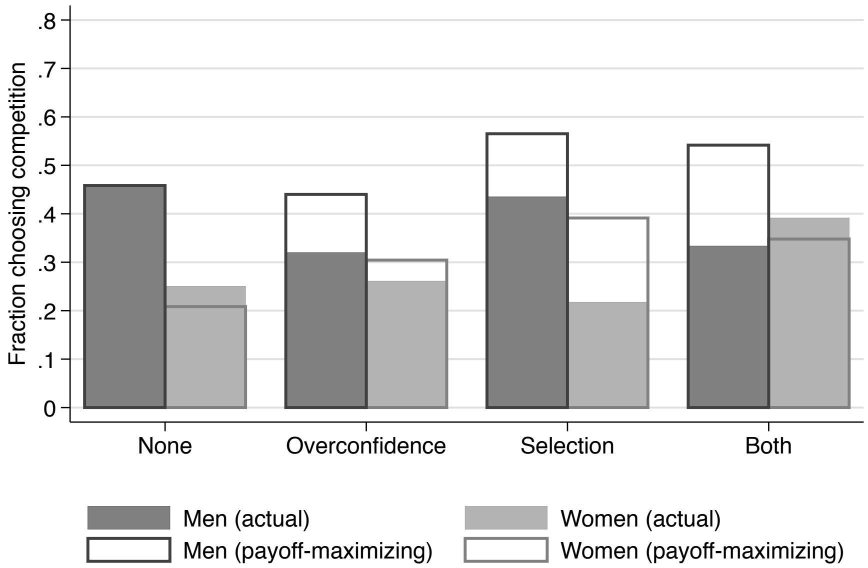
that is even stronger than for previous performance. Interestingly, however, there is a small percentage of women who enter the competition even though they think that their probability of winning is below 25%.

The fraction of men and women selecting the competitive incentive scheme by treatment is depicted in Figure 13. Apart from the fact that absolute ratios of subjects deciding to compete seem pretty low when compared to previous literature, it can also be seen that while more men than women decide to compete without feedback, gender differences seem to be ameliorated when feedback is available in round three. Competing against a selected or a random sample does not seem to make a difference to men and to women only if combined with feedback; thus, subjects appear to exhibit competition neglect. Testing for differences between genders within treatment and for differences between the treatments with Fisher exact tests does not detect significant differences, however.[22]

To explore whether there are differential effects of the treatments for men and women, I build on the regression specification shown in Table 3 and include interactions of the treatment conditions with gender. The results are shown in Table 6. It can be seen that while the main effect of gender still is insignificant and previous performance and beliefs still have significant influence, there is an additional effect in the full model: when

[22] Note that while the sample size does not allow to detect differences in this experiment, the number of subjects would have been sufficient to detect gender differences based on previous results. As ex ante power calculations were based on competition ratios in existing experiments, this was unexpected.

Figure 13: *Fraction of individuals selecting into competition by treatment*



Notes: Figure depicts fraction of men and women choosing the competitive incentive scheme in round three of the main sessions. Calculations are based on 96 men and 93 women.

controlling for beliefs and positive performance, the interaction between the feedback indicator and gender is significant. Hence, providing feedback increases the probability that women select the competitive incentive scheme.

D. Analysis of Round 4: Competition Based on Past Performance

In the experiment, an additional round was included after the three rounds described in the paper, the difference being here that in round three, participants made the choice on their subsequent performance while in round four, they were again paid based on their performance in round two. This was done to address the fact that factors like differing attitudes against the act of competing or ambiguity attitudes are discussed to influence competition decisions (see Niederle and Vesterlund, 2011).

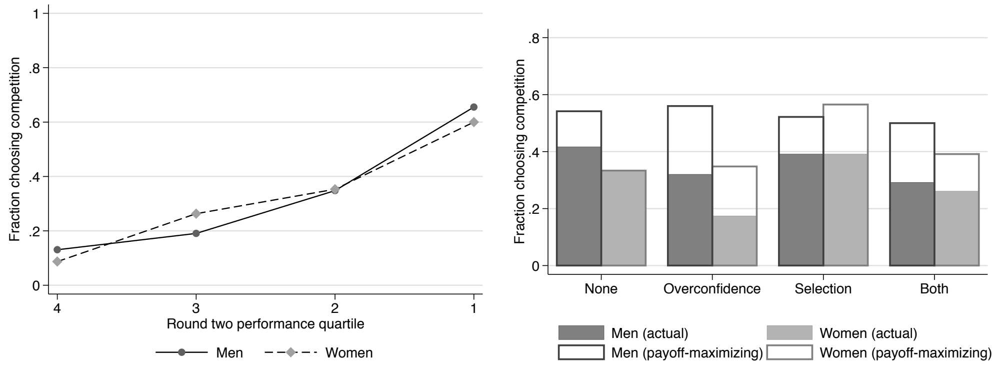
First, there is even less of a gender difference in competition entry than in round three: while 35.42% of men chose the competitive incentive scheme, 29.03% of women do. Considering actual performance in the reference group, a payoff-maximizing individual should choose competition if her round-two performance is at least seven without and eight with selection in the competition group. Based on round two performances, this applies to 46.81% of men and 35.42% of women.

Table 6: *Determinants of competition choices*

DV: Competition (Y/N)	(1)	(2)	(3)	(4)
Feedback = 1	-0.106	-0.0950	-0.113	-0.0965
	(0.0944)	(0.0937)	(0.0860)	(0.0843)
Selection = 1	-0.00256	-0.0198	0.0308	0.0200
	(0.0929)	(0.0991)	(0.0845)	(0.0875)
Feedback # Female	0.196	0.208	0.240**	0.223*
	(0.127)	(0.132)	(0.112)	(0.117)
Selection # Female	-0.100	-0.0803	-0.124	-0.114
	(0.127)	(0.139)	(0.115)	(0.125)
Previous performance	0.0581***	0.0602***	-0.0262	-0.0256
	(0.0105)	(0.0113)	(0.0236)	(0.0235)
Own performance belief			0.110***	0.115***
			(0.0317)	(0.0323)
Others' performance belief			-0.0870***	-0.0866***
			(0.0273)	(0.0297)
Subj. winning probability			0.621***	0.602***
			(0.168)	(0.175)
Female = 1	-0.0820	-0.0692	-0.0278	-0.00136
	(0.115)	(0.129)	(0.0971)	(0.107)
Risk aversion	-0.0449**	-0.0434**	-0.0316*	-0.0307*
	(0.0178)	(0.0179)	(0.0165)	(0.0170)
Controls	No	Yes	No	Yes
Observations	189	187	189	187
Adjusted R^2	0.158	0.156	0.320	0.314

Notes: Table reports results of linear probability estimations. Robust standard errors in parentheses. Estimations are based on round three. Controls include age, an indicator for field of study (business, economics, STEM, humanities, law, psychology, other social sciences, other), and math grade. *** p<0.01, ** p<0.05, * p<0.1.

Figure 14: *Competition choices by performance and treatment in round four*



Notes: Figure depicts fraction of men and women choosing the competitive incentive scheme in round four of the main sessions by performance quartile (left panel) and by treatment (right panel). Calculations based on 96 men and 93 women.

Exploring competition decisions based on actual performance and by treatment in Figure 14, it can be seen that as in round three, there is a positive relationship between performance quartile and proportion of individuals selecting into competition that looks similar for men and women. With respect to treatment variations, it can be seen that there are roughly similar entry ratios for men facing the random and the selected sample of competitors, while entry ratios for women are lower when they have to compete against competitors who themselves self-selected into the competitive incentive scheme. Feedback is associated with lower entry rates for both men and women.

E. Instructions

General Introduction

Welcome to the experiment and thank you for your participation!

Please do not speak with the other participants from now on. Please remain silent throughout the course of the experiment.

This experiment is designed to evaluate economic decision-making behavior. You will be able to earn money which will be paid to you after the experiment privately and in cash.

During the experiment, you will be asked to make several decisions. Some of them will be made in interaction with other participants. This means that both your own decisions and those of the other participants may determine your payoffs.

The entire experiment will last for about 60 minutes and consists of two parts. At the beginning of each part you will receive detailed instructions. The parts of the experiment are independent of each other, i.e., decisions in one part do not have any impact on your earnings in the other parts. The sum of your earnings from all parts will be added to your total earnings in this experiment. The total earnings will be paid to you upon completion of the final part, individually and in cash.

Payoff During the course of the experiment, payoffs are calculated not in Euros, but in Experiment Points (EP). At the end of the experiment, the sum of your earned EP will be converted to Euros. Here, the following exchange rate applies: 100 EP = 1 €.

In addition to the income that you can earn during the experiment, you receive 5 € for your punctuality.

Anonymity None of the other participants will be able to observe your choices in this experiment. In addition, the data from the experiment will be evaluated anonymously. At the end of the experiment, you have to sign a receipt for the income you have earned during this experiment. This is only due to accounting issues and cannot be used to associate your personal information with your decisions. Your name cannot be combined with your behavior in the experiment at any time.

Permitted aids There are a pen and notepaper located on your desk. Please leave them on the table after the experiment.

If you have any questions after the instructions or during the experiment, please press the red button on the keyboard (F11). One of the experimenters will then answer your question in private. If you do not need help any more, please press the red button again.

Introduction of Part 1 Procedure

In this part of the experiment, you will go through four rounds. In every round, you will have to work on two different tasks: a sum-calculation task and an estimation task.

Sum calculation In this task, you are requested to calculate sums of five random two-digit numbers. You will have a time window of three minutes for calculating as many sums as possible. It is not allowed to use a calculator, but you can use the provided notepaper.

Your screen is built up as follows: On the left side, you are going to see the five two-digit numbers you should add up. On the right, there is a text box where you enter the solution. Then, you press “OK” (on the far right) to submit your answer and going to the next calculation. On the top right side you can check the time you have left in this round.

– Participants have one minute to get accustomed to the screen for the sum calculations –

Payment rules You are paid according to different rules in every round. At the beginning of each round, you are informed about how your earning is determined.

In some rounds, your payment can also depend on the performance of other participants. In this case, this does not apply to the performance of the other participants sitting in the lab with you right now, but participants of a past experiment. These were selected in the same way as you, have worked on the same tasks and made the same decisions as you. Therefore, these participants are comparable to you and the other participants of this session. In the following, whenever you read about the "reference group", these comparable participants of an earlier session are meant.

Estimation task In this task you are requested to give your own estimation. This will be explained more in detail in the following rounds.

Summary of rounds In summary, all rounds are going to follow the same procedure:

- You are informed about the payoff rule for this round
- You do the estimation task
- You have three minutes to finish the sum calculations

At the end of the experiment, one out of the four rounds will be chosen randomly and your earnings from that round paid together with the 5 € show-up fee and your payoff from Part 2.

Explanation of Estimation Task

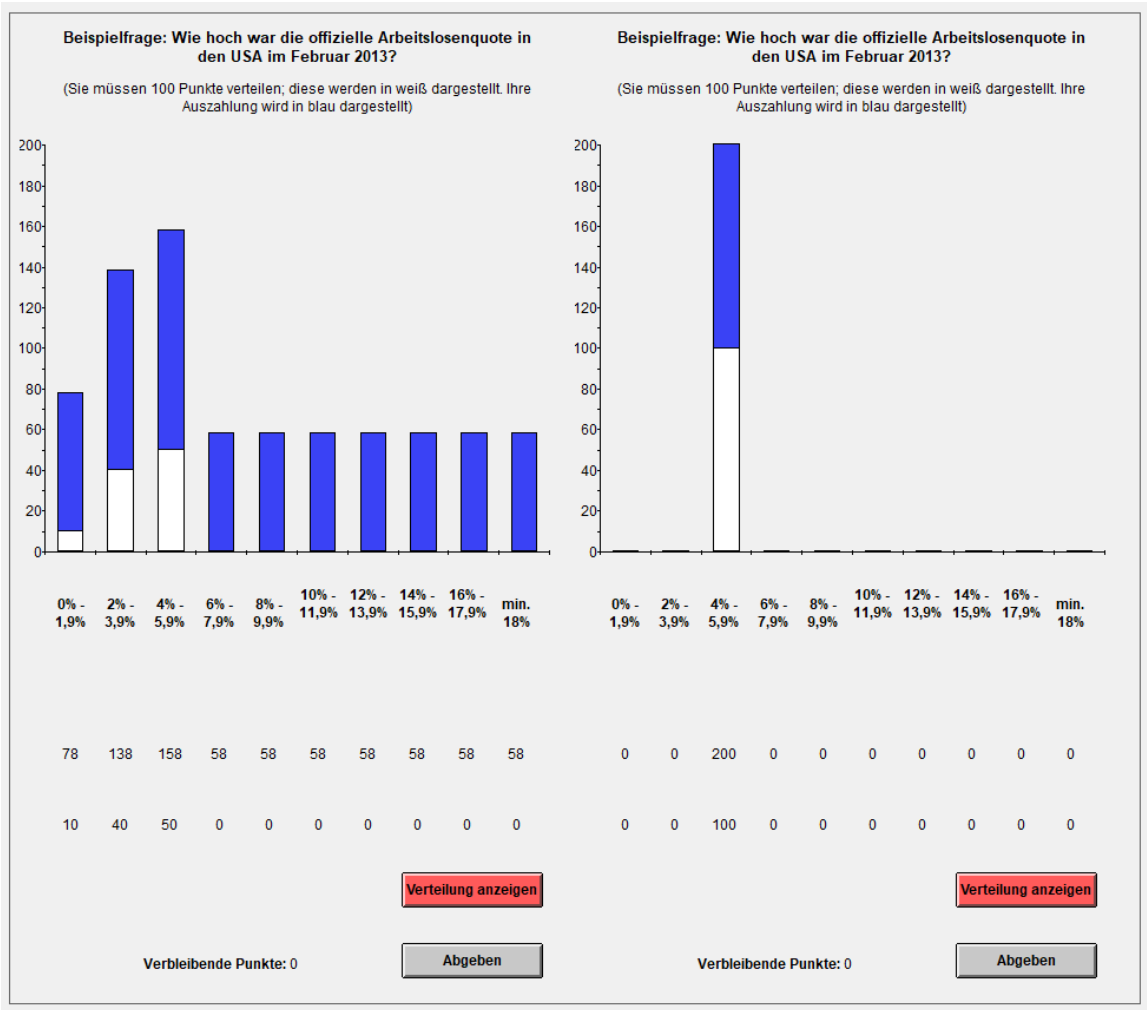
Estimation Task 1 In the following you are asked two questions. You are asked to state what you think the correct answer to the respective question is.

For each question, you can allocate a total of 100 points to 10 possible categories. To do so, select the check box below the particular category and type in how many of your 100 available points you want to allocate to that category. By pressing "show distribution" you can see your chosen distribution of points between the categories looks like (bars on the upper part of the screen). You have to allocate the total amount of 100 points available to you. Once you are satisfied with the distribution, press "submit" to get to the next screen.

In total, you can earn up to 200 EP in each question. For this, the points you have distributed are converted into EP in consideration of the full distribution. The EP you earn when you submit the currently chosen distribution are displayed to you as as bars and in numbers above the check boxes as soon as you click on "show distribution".

How much points you should allocate to each category depends on your estimation of the correct answer to the question asked. To illustrate this point, consider the following example:

How high was the official unemployment rate in the US in February 2003?



You can now freely choose how you want to spread your 100 points. The figure shows you two exemplary options.

- *Left distribution:* Assume you think it very likely that the true answer is slightly under 5%. Then you could allocate 50 points to the category "4 to 5.9%", 40 points to the category "2% to 3.9%", and the remaining 10 points to the category "0-1.9%", for example.
  Your payoff for this question is determined by the amount of points you have allocated to the category that contains the true unemployment rate. If it is actually between 4 and 5.9%, you earn 158 EP. If it is between 2 and 3.9%, you receive 138 EP, and if it is between 0 and 1.9%, you receive 78 EP. For each other true unemployment rate you receive with this distribution 58 EP. Thus, your finally payoff depends on your given estimation and the correct answer to the questions.

You can change the allocation of your points as much as you want to such that it represents your personal estimation as well as possible. As the actual unemployment rate in the USA in February 2013 was 7.7%, you would have earned 58 EP with the left distribution.

- *Right distribution:* Assume you had allocated all your available points on one category, for example on the category “4% to 5.9%”. Then, the distribution of possible earnings for this question would look as follows: If the true unemployment rate lies between 4 and 5.9%, you would earn the maximum payoff of 200 EP for this question. However, since the true unemployment rate is 7.7%, you would not have earned anything with this distribution of points.

How you find a trade-off between the precision of your estimation and the risk that you are wrong is up to you. Keep in mind the following three important things:

- Your assessment of the right answer to the question is a personal estimation which is based on the information you have.
- Depending on your decision you can earn up to 200 EP per question.
- Your decisions can also depend on your willingness to take risks. The estimations you are going to make now are individual.

Estimation Task 2: Winning Probability Now we would like you to give an estimation of how likely it is that you are going to solve more calculations correctly that a randomly chosen participant from the reference group.

As a reminder: The reference groups consist of participants of a past date which have solved the same tasks you will solve now.

Your earnings from this estimation are affected by the precision with which you specify the probability that your performance is above the performance of a randomly chosen participant from the reference group. The payoff is constructed such that you have the highest chance of earning money if you indicate your actual assessment. The mechanism works as follows: You indicate your estimated winning probability. In addition, the computer draws a number X between 0 and 100. Every number between 0 and 100 is drawn with the same probability.

A comparison of these two values determines according to which of the following criteria you are paid for your estimation.

- Option A: you receive 200 EP if you solve more calculations correctly than a randomly chosen participant of the reference group

- Option B: you receive 200 EP with probability X

The option according to which you are paid is selected as follows: If X is higher than your indicated winning probability, option B is chosen and you receive 200 EP with probability X. If X is lower than your indicated winning probability, option A is selected and you receive 200 EP if your performance in this round is better than the one of a randomly chosen subject of the reference group.

This means that you maximize your chance of winning the price of 200 EP by indicating your actual estimation of how likely you think it is to be better than a randomly chosen subject of the reference group.

An example: You think that you can solve more calculation tasks correctly that a randomly chosen person of the reference group with a probability of 62%.

- If you truly state this estimation, you are paid according to option A if X is lower than 62, which gives you a higher winning probability (namely 62%) than option B (X%). If X is higher than 62, you are paid according to option B, which gives you a higher chance of winning (namely X%) that option A (62%). Regardless of which number X is drawn randomly, you are always paid according to the rule that gives you a higher chance of receiving 200 EP.
- If you state another probability, for example 10%, you are paid according to option A if X is smaller than 10, which gives you a higher winning probability (namely 62%) than option B. If X is higher than 10, you are paid according to option B, which possibly gives you a lower winning probability than option A. If X is between 10 and 62, you are paid according to option B, although with option A you would have had a higher winning probability of 62%. In this case, stating a false probability reduces your chance to receive 200 EP.

The logic of this example applies to all other probabilities as well. Whatever you think how likely it is that you are going to solve more calculations correctly than a randomly chosen subject from the reference group, you maximize the chance of receiving 200 EP by giving your best possible estimation.

Payment Descriptions

Round 1 You receive 50 EP for each correct calculation.

Round 2 The performance of an individual from the reference group in this round is chosen randomly. Remember: the reference group consist of participants of a past date who solved the same tasks you are solving now. If you solve more calculations correctly than this person, you receive 100 EP per correct calculation; otherwise, you do not get any payoff for this round. If you solve the same number of calculations correctly as the randomly selected reference person, the computer will randomly choose a winner.

Round 3 In this round, you can choose how you want to be paid.

- *Option A*: As in round 1, you receive 50 EP for each correct calculation.
- *Option B*: As in round 2, the performance of an individual from the reference group in this round is chosen randomly. If you solve more calculations correctly than this person, you receive 100 EP per correct calculation; otherwise, you do not get any payoff for this round. If you solve the same number of calculations correctly as the randomly selected reference person, the computer will randomly choose a winner.

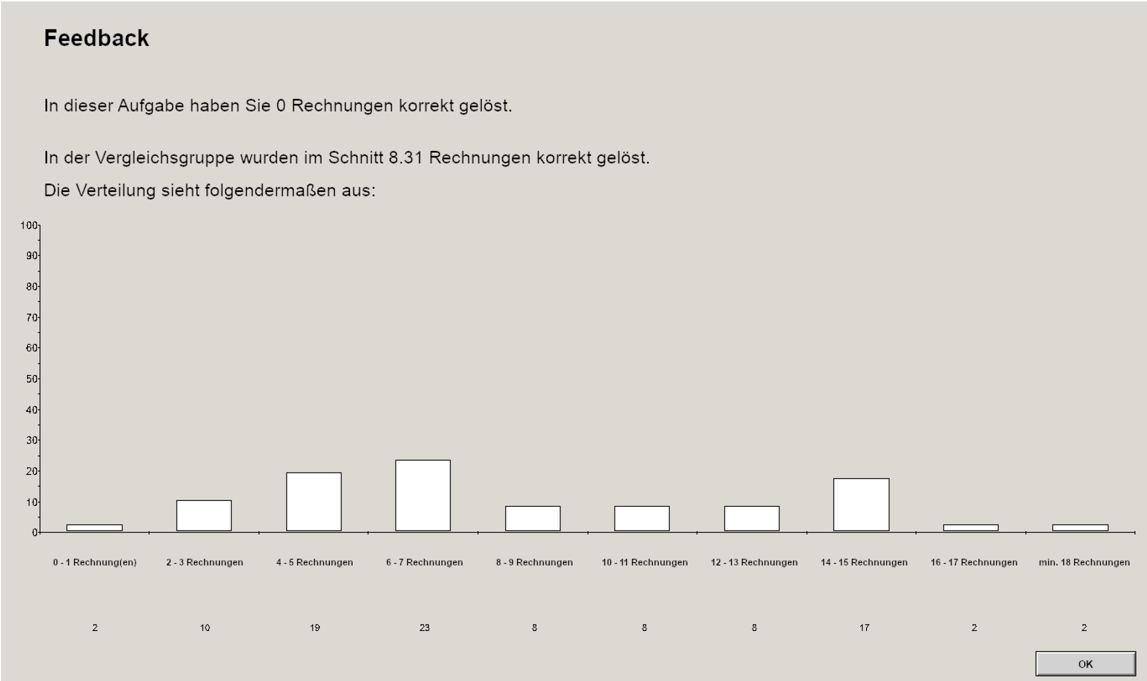
If the competitor was drawn from the selected competition group, the following sentence was added to Option B: In contrast to round 2, the person is selected only among the other participants in the reference group who have chosen option B.

Round 4 In this round, the number of correct calculations will again determine your payoff. However, in contrast to the previous rounds, you are not going to work again on the task. Instead, your performance in round 2 determines your earnings. You can choose how you want to be paid.

- *Option A*: As in round 1, you receive 50 EP for each correct calculation.
- *Option B*: As in round 2, the performance of an individual from the reference group in this round is chosen randomly. If you solve more calculations correctly than this person, you receive 100 EP per correct calculation; otherwise, you do not get any payoff for this round. If you solve the same number of calculations correctly as the randomly selected reference person, the computer will randomly choose a winner.

If the competitor was drawn from the selected competition group, the following sentence was added to Option B: In contrast to round 2, the person is selected only among the other participants in the reference group who have chosen option B.

Feedback



This translates to: "In this task, you solved X sum correctly.

In the reference group, an average number of Z calculations were solved correctly. The distribution of performances looks as follows"

References

Arkes, H. R., Christensen, C., Lai, C., and Blumer, C. (1987). Two methods of reducing overconfidence. *Organizational Behavior and Human Decision Processes*, 39(1):133–144.

Balafoutas, L., Kerschbamer, R., and Sutter, M. (2012). Distributional preferences and competitive behavior. *Journal of Economic Behavior & Organization*, 83(1):125–135.

Bartling, B., Fehr, E., Maréchal, M. A., and Schunk, D. (2009). Egalitarianism and competitiveness. *American Economic Review*, 99(2):93–98.

Benoît, J.-P. and Dubra, J. (2011). Apparent overconfidence. *Econometrica*, 79(5):1591–1625.

Berlin, N. and Dargnies, M.-P. (2016). Gender differences in reactions to feedback and willingness to compete. *Journal of Economic Behavior & Organization*, 130:320–336.

Bolger, F., Pulford, B. D., and Colman, A. M. (2008). Market entry decisions: Effects of absolute and relative confidence. *Experimental Psychology*, 55(2):113–120.

Buser, T., Gerhards, L., and Van Der Weele, J. J. (2016). Measuring responsiveness to feedback as a personal trait. Working Paper.

Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, competitiveness, and career choices. *Quarterly Journal of Economics*, 129(3):1409–1447.

Cain, D. M., Moore, D. A., and Haran, U. (2015). Making sense of overconfidence in market entry. *Strategic Management Journal*, 13(1):1–18.

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

Dechenaux, E., Kovenock, D., and Sheremeta, R. M. (2015). A survey of experimental research on contests, all-pay auctions and tournaments. *Experimental Economics*, 18(4):609–669.

Dohmen, T. and Falk, A. (2011). Performance pay and multidimensional sorting: Productivity, preferences, and gender. *American Economic Review*, 101(2):556–590.

Ducharme, W. M. and Donnell, M. L. (1973). Intrasubject comparison of four response modes for "subjective probability" assessment. *Organizational Behavior and Human Performance*, 10(1):108–117.

Eckel, C. C. and Grossman, P. J. (2008). Men, women and risk aversion: Experimental evidence. In Plott, C. R. and Smith, V. L., editors, *Handbook of Experimental Economics Results*, volume 1, pages 1061–1073. Elsevier.

Enke, B. and Zimmermann, F. (2017). Correlation neglect in belief formation. Working Paper.

Eriksson, T., Teyssier, S., and Villeval, M.-C. (2009). Self-selection and the efficiency of tournaments. *Economic Inquiry*, 47(3):530–548.

Esponda, I. (2008). Behavioral equilibrium in economies with adverse selection. *American Economic Review*, 98(4):1269–1291.

Evans, J. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59:255–278.

Evans, J. and Stanovich, K. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8(3):223–241.

Eyster, E. and Rabin, M. (2005). Cursed equilibrium. *Econometrica*, 73(5):1623–1672.

Flory, J. A., Leibbrandt, A., and List, J. A. (2015). Do competitive workplaces deter female workers? A large-scale natural field experiment on job-entry decisions. *Review of Economic Studies*, 82(1):122–155.

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in competitive environments: Gender differences. *Quarterly Journal of Economics*, 118(3):1049–1074.

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1):114–125.

Grossman, S. J. and Hart, O. D. (1983). An analysis of the principal-agent problem. *Econometrica*, 51(1):7–45.

Grossman, Z. and Owens, D. (2012). An unlucky feeling: Overconfidence and noisy feedback. *Journal of Economic Behavior & Organization*, 84(2):510–524.

Harrison, G. W., Martínez-Correa, J., Swarthout, J. T., and Ulm, E. R. (2015). Eliciting subjective probability distributions with binary lotteries. *Economics Letters*, 127:68–71.

Harrison, G. W., Martínez-Correa, J., Swarthout, J. T., and Ulm, E. R. (2017). Scoring rules for subjective probability distributions. *Journal of Economic Behavior & Organization*, 134:430–448.

Hölmstrom, B. (1979). Moral hazard and observability. *Bell Journal of Economics*, 10(1):74–91.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5):1644–1655.

Jehiel, P. (2016). Investment strategy and selection bias: An equilibrium perspective on overoptimism. Working Paper.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kamas, L. and Preston, A. (2012). The importance of being confident; gender, career choice, and willingness to compete. *Journal of Economic Behavior & Organization*, 83(1):82–97.

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

Larkin, I. and Leider, S. (2012). Incentive schemes, sorting, and behavioral biases of employees: experimental evidence. *American Economic Journal: Microeconomics*, 4(2):184–214.

Lazear, E. P. (2000). Performance pay and productivity. *American Economic Review*, 90(5):1346–1361.

Merkle, C. and Weber, M. (2011). True overconfidence: The inability of rational information processing to account for apparent overconfidence. *Organizational Behavior and Human Decision Processes*, 116(2):262–271.

Möbius, M. M., Niederle, M., Niehaus, P., and Rosenblat, T. S. (2014). Managing self-confidence. Working Paper.

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

Moore, D. A., Oesch, J. M., and Zietsma, C. (2007). What competition? Myopic self-focus in market-entry decisions. *Organization Science*, 18(3):440–454.

Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, 122(3):1067–1101.

Niederle, M. and Vesterlund, L. (2011). Gender and competition. *Annual Review of Economics*, 3(1):601–630.

Schlag, K. H., Tremewan, J., and Van der Weele, J. J. (2015). A penny for your thoughts: A survey of methods for eliciting beliefs. *Experimental Economics*, 18(3):457–490.

Simonsohn, U. (2010). eBay's crowded evenings: Competition neglect in market entry decisions. *Management Science*, 56(7):1060–1073.

van Veldhuizen, R. (2017). Gender differences in tournament choices: Risk preferences, overconfidence or competitiveness? Working Paper.

Vandegrift, D., Yavas, A., and Brown, P. M. (2007). Incentive effects and overcrowding in tournaments: An experimental analysis. *Experimental Economics*, 10(4):345–368.

Wozniak, D., Harbaugh, W. T., and Mayr, U. (2014). The menstrual cycle and performance feedback alter gender differences in competitive choices. *Journal of Labor Economics*, 32(1):161–198.