Strategic Competition and Self-Confidence

Working Paper

Author(s):
Brilon, Stefanie; Grassi, Simona; Grieder, Manuel; Schulz, Jonathan

Publication date:
2019-11-14

Permanent link:
https://doi.org/10.3929/ethz-b-000392522

Originally published in:
SSRN 19-40, https://doi.org/10.2139/ssrn.3481057

DEPARTMENT OF ECONOMICS

STRATEGIC COMPETITION AND SELF-CONFIDENCE

STEFANIE BRILON
SIMONA GRASSI
MANUEL GRIEDER
JONATHAN F. SCHULZ

George Mason University
Department of Economics
Working Paper No. 19-40

Strategic Competition and Self-Confidence

Stefanie Brilon* Simona Grassi† Manuel Grieder‡§
Jonathan F. Schulz¶

October 29, 2019

Abstract

We test the hypothesis that strategic interactions foster overconfidence. We experimentally compare an environment where players have an incentive to overstate their own ability to deter competitors, with one where this incentive is removed. We find that overconfidence persists in the former environment but vanishes in the latter. The persistence of overconfidence in the strategic environment is driven by three mechanisms. First, players who win uncontested update their confidence as if they had won in actual competition. Second, in contrast, individuals who do not compete do not update their confidence, thus creating an asymmetry in updating. Third, inflated confidence signals of ability are "contagious" because they affect how their receivers update their confidence. We provide empirical evidence that these mechanisms can explain stylized facts on overconfidence such as the Dunning–Kruger effect. We also discuss implications for organizational design and management.

Keywords: Overconfidence, competition, strategic deterrence, self-deception, belief updating

JEL Codes: A12, C91, D83, D90, D91

*Philipps Universität Marburg, Department of Economics, stefanie.brilon@wiwi.uni-marburg.de

†King's Business School, Economics, King's College London, simona.grassi@kcl.ac.uk

‡ETH Zurich, Chair of Economics, manuel.grieder@econ.gess.ethz.ch

§Zurich University of Applied Sciences (ZHAW), School of Management and Law, manuel.grieder@zhaw.ch

¶George Mason University, Department of Economics, and CeDEx, jonathan.schulz77@gmail.com

For helpful comments we thank Simon Gächter, Lorenz Goette, Luís Santos-Pinto, and Christian Zehnder as well as seminar audiences in Cape Town, Frankfurt, Lausanne, Nottingham, Vienna, and Zurich (UZH & ETH), and conference participants in Bern (BEER), Florence (BEELab workshop), Toronto (BDRM), and Rome (IMEBESS).

1 Introduction

Evidence from psychology and economics suggests that many people hold exaggerated beliefs about their own abilities compared with those of others (see, e.g., Dunning et al., 1989; Merkle & Weber, 2011; Burks et al., 2013; Benoît et al., 2015). Such overconfidence can be costly for individuals, organizations, and society as a whole. It has, for instance, been invoked as an explanation for excessive entry into markets (Camerer & Lovallo, 1999), excessive trading and stock market crashes (Barber & Odean, 2001; Scheinkman & Xiong, 2003), value-destroying corporate mergers (Malmendier & Tate, 2008), and even wars (Johnson, 2009).

How can individual overconfidence persist even if it leads to decisions based on incorrect presumptions and, often, costly mistakes? In this paper, using a laboratory experiment, we test the hypothesis that overconfidence is persistent in competitive environments where players have incentives to strategically display signals of strength and dominance to deter opponents.

In social contexts, competition is ubiquitous. In organizations, individuals face competition in the form of job searches or evaluations for promotion. In other domains of life, competitive situations range from mating to sports to political elections. Typically, one can choose whether or not to enter a competition, with the decision being likely to depend on how one assesses her own ability in comparison with the ability of opponent(s). In such contexts, it may be advantageous to behave strategically and send signals of strength to deter opponent(s) from competing. Thus, there exists a potential benefit of appearing more confident. In this study, we demonstrate and explain why the option to opt into or out of competition, and the possibility of deterring opponents with messages about one's own strength, sustain overconfidence in the long run. We also show that, instead, in a competitive environment where one is forced to compete, the accumulated feedback helps to provide a better assessment of where one stands relative to others, and overconfidence does not persist.

We designed an incentivized laboratory experiment where participants are exposed to either a strategic or a non-strategic competitive environment. In the *strategic* treatment, participants repeatedly decide to enter or stay out of competition after having received ability signals from a prospective opponent. This creates the strategic possibilities of bluffing and deterring opponents and of being deterred. By contrast, in our *non-strategic* baseline treatment, players cannot send ability signals and are forced to compete. There is thus no room for strategic bluffing or deterrence in this condition. Subsequently, we measure the impact of these two environments on confidence. We find that overconfidence persists in the strategic environment, whereas it decreases and vanishes in the non-strategic environment.

In further treatments (based on sender–receiver games), we demonstrate that at least three factors contribute to sustained overconfidence in a strategic competitive environment. First, we find that successfully deterring an opponent has the same impact on confidence as winning in an actual competition. This has the striking implication that successful bluffers, i.e., players who deter an opponent who would have defeated them in actual competition, become more confident. Second, players who choose to stay out of competition do not on average update their confidence. Since they do not obtain feedback on the actual outcome of a competition, this is unsurprising. However, it is in contrast to their opponents who do update as if they had won. This asymmetry in updating fosters overconfidence on average in the population. Third, we find that inflated ability signals have a "contagious" effect on confidence; i.e., individuals base their confidence updates not only on the information about whether they have won or lost, but also on the signals they received from their opponents. This is particularly the case for the losing players: the higher the ability signal they receive from their opponent, the less they decrease their confidence.

These findings demonstrate that the experiences accumulated in social strategic environments shape self-confidence. We further show that the mechanisms we uncover provide a novel explanation for two well-known empirical patterns, namely, the Dunning–Kruger effect (i.e., that individuals with relatively low skill levels are more overconfident) and that women are less confident than men.

Kruger & Dunning (1999) argue that low-skilled individuals lack the metacognitive ability to correctly judge their relative ability and are therefore more likely to be overconfident. In our study, we also find that after being exposed to a strategic environment, relatively low-skilled participants are overly optimistic about their relative ability. We provide evidence that this is not due to lower metacognitive skills, but instead depends on the different experiences of low- and high-skilled individuals in the strategic competitive environment, where the low-skilled individuals are less likely to enter into competition and hence receive less feedback about their relative ability. Thus, they are less likely to adjust their confidence downward.

Similarly, the well-documented gender differences in confidence can be explained by different experiences in competitive environments for men and women (Bertrand, 2011). Empirical work has established that women are more honest than men. Furthermore, women are found to be more reluctant to compete than men (see, e.g., Croson & Gneezy, 2009; Dreber & Johannesson, 2008; Gneezy et al., 2003; Gneezy & Rustichini, 2004; Niederle & Vesterlund, 2007). In our work, we find evidence that women are less likely than men to bluff and win uncontested when they would otherwise have lost. We find that women also avoid competition more often than men, thus missing opportunities to experience wins and positively update their confidence. We conclude that different attitudes toward ubiquitous strategic competitions in organizations and other domains of life can perpetuate the persistence of differences in confidence levels between genders.

Our study offers novel insights to explain why individuals who hold coveted positions in society, such as government leaders and CEOs, display high self-confidence—often indistinguishable from overconfidence (see, e.g., Goel & Thakor, 2008; Hiller & Hambrick, 2005). Confidence could have helped them to reach those positions by facilitating the deterrence of opponents in the first place, and, as our results show, deterring opponents increases confidence, a mechanism that feeds on itself. Besides, once one is in a top position, the chances of being contested by subordinates gets even smaller.

Leaders' personal attributes have an important impact on organizational performance (see, e.g., Hambrick & Mason, 1984; Hambrick, 2007; Nadkarni & Herrmann, 2010; Peterson et al., 2003; Wang et al., 2016). A vast literature in finance and management has stressed the detrimental effects of managers' overconfidence on firm performance (see, e.g., Camerer & Lovallo, 1999; Barber & Odean, 2001; Malmendier & Tate, 2005, 2008; Banerjee et al., 2018).[1] However, overconfident managers can also potentially enhance a firm's value, for instance, by engaging in ambitious projects (see, e.g., Bénabou & Tirole, 2002; Galasso & Simcoe, 2011; Hirshleifer et al., 2012; Phua et al., 2018).[2] In our study, we do not take a stand on whether leader overconfidence is detrimental or beneficial to organizations. Rather, our results offer an insight into optimal organizational design, namely, that the level and the persistence of overconfidence within an organization can be determined by a conscious design of its competitive structures (e.g., by encouraging competition across the workforce and/or implementing appropriate feedback systems.)

Our study is related to an emerging literature that emphasizes the role played in social signaling by overconfidence (Burks et al., 2013; Charness et al., 2018; Ewers & Zimmermann, 2015; Proeger & Meub, 2014; Schwardmann & van der Weele, 2019; Trivers, 2011).[3] Trivers (2011) explains how natural selection favors self-deception since self-deception facilitates interpersonal deception. Burks et al. (2013) provide evidence that overconfidence is induced by the desire to send positive signals to others about one's own skill, more than by the need to preserve a positive self-image. Ewers & Zimmermann (2015) show that individuals become

[1]An example is Malmendier & Tate (2005), who show how overconfident managers are prone to making suboptimal investment decisions, since they overestimate the return to their investment projects.

[2]An example is Galasso & Simcoe (2011), who present evidence that for large-firm CEOs, overconfidence is associated with a greater propensity to innovate.

[3]There have been a number of analyses of the role played by overconfidence outside social contexts. Overconfidence can be explained by its consumption value (see, e.g., Brunnermeier & Parker, 2005; Kőszegi, 2006; Weinberg, 2009) and its effect on motivating effort (see, e.g., Bénabou & Tirole, 2002; Compte & Postlewaite, 2004). Moreover, overconfident beliefs may emerge as a result of rational Bayesian updating under certain information structures (Benoît & Dubra, 2011).

more overconfident when confidence is publicly observable. Schwardmann & van der Weele (2019) provide intriguing empirical evidence that more overconfident individuals are more persuasive in convincing others of their superior performance and hence that overconfidence pays off. We provide evidence supporting the social signaling motivation for overconfidence: overconfidence is only sustained in a competitive strategic environment, in which ability signals are important as a means to facilitate interpersonal deception. A non-strategic environment does not foster overconfidence. The psychological mechanism (namely, biased updating after winning solely by deterrence) that fosters overconfidence in the strategic environment may ultimately foster self-deception, which in turn helps to more effectively deter others. While the literature on social signaling has contrasted social motives or self-image concerns as drivers of overconfidence, we find that social signaling and self-image concerns are not mutually exclusive drivers of overconfidence, but can coexist. The excessive reliance on received ability signals after losing (the "contagious" effect of confidence) can certainly be conducive to preserving a positive self-image.

Our experimental design builds on Charness et al. (2018). They demonstrate that in a one-shot sender–receiver game, participants inflate their stated confidence when it is strategically beneficial for them to do so. We extend their design to multiple repeated interactions and measure confidence also when strategic incentives are removed. We show that strategic behaviors in a competitive environment foster overconfidence that persists even when the strategic incentives are no longer present.

Our major contribution is to show that overconfidence is shaped by repeated human social interactions—here strategic competitive environments. Biased updating and persistent overconfidence emerge as a result of experiences shaped by incentives in social environments. Our focus on the role played by accumulated experiences in specific environments relates our study to the recent literature that emphasizes variation in confidence across societies, domains, or socio-economic status (see, among others, Muthukrishna et al., 2018; Belmi et al., 2019). These differences likely reflect variations in accumulated experiences across those different environments.

The remainder of this paper is organized as follows. Section 2 presents the experimental design. The results are presented in Sections 3 and 4. Section 5 concludes and discusses the implications of our findings. Appendix A contains robustness analyses, Appendix B contains the experimental instructions (translated from French), and Appendix C provides a theoretical model to back up our experiment.

2 Design

To investigate the impact of a strategic environment on confidence, we randomly allocated participants either to a non-strategic or a strategic environment and subsequently elicited their confidence. In the non-strategic baseline, we ruled out strategic incentives by forcing players to compete with their opponents. In the strategic treatment, players exchanged ability signals and, upon receiving their opponent's signal, they could decide whether to enter the competition or not.

Both of these treatments are *symmetric* in nature; that is, both players have the same actions available. While this may be a more realistic feature of these treatments, it makes it more difficult to disentangle the effects of strategic behavior on confidence in our experiment. We therefore conducted two *asymmetric treatments*, where only one player (the Sender) can send an ability signal, whereas only the other player (the Receiver) can opt out of the competition. These asymmetric treatments help us to further investigate the mechanisms that foster overconfidence in the strategic environment and are described in more detail below. In all treatments, the experiment consisted of four stages: Stage 1 determined the individuals' IQ scores; Stage 2 elicited the participants' initial confidence; Stage 3 implemented the treatment variation; and Stage 4 elicited the final confidence after individuals experienced the different environments. A detailed description of each stage and treatment variation follows.

Stage 1: IQ. In Stage 1, all participants took an eight-minute-long IQ test consisting of 15 Raven Advanced Progressive Matrices (Raven, 2000). We used the same 15 questions as Charness et al. (2018). The number of correct answers (i.e., the participants' "IQ score") was never disclosed during the experiment. Participants were told that a high IQ score was likely to lead to higher earnings in subsequent stages, without further details.

Stage 2: Initial Confidence. After the IQ test, we asked participants how likely it was that their IQ score placed them in the top half of a group of ten randomly assigned participants (see Charness et al., 2018; Moebius et al., 2014; Zimmermann, 2019). Participants stayed in the same group of ten during the entire experiment and were informed about it.

We used a *quadratic scoring rule* to incentivize accurate and truthful confidence elicitation: if a participant reported to be in the top half of her group with probability p, she received a payoff (in experimental currency) of $\pi_{\text{top}} = [1 - (1-p)^2] \times 10$ if she was indeed in the top half and of $\pi_{\text{bottom}} = (1 - p^2) \times 10$ if she was not. If a participant was in the top half, they would gain the highest number of points (10) by reporting $p = 1$. Similarly, if a participant was in the lower half, the highest number of points would be gained by stating

$p = 0$. Incorrect reporting was penalized with increasing decrements from 10 the further the participants were from 1 or 0. In the instructions to participants, we explained with several examples that this rule penalized both an inaccurate and an insincere report of confidence (our explanations closely followed those of Charness et al. (2018).)

Stage 3: Competition and Treatments. In Stage 3, we implemented our treatment variation. All treatments shared a common underlying structure. Each treatment consisted of 12 rounds, with each round consisting of three steps. In the first step, each participant was randomly paired with one anonymous opponent out of her previously determined group of ten participants. In the second step, all participants were asked to report their relative ability (i.e., the probability of being in the top half of their reference group). Again, this elicitation was incentivized according to the quadratic scoring rule.

In the third step, competition took place; that is, the IQ scores of each pair were compared. The treatment variation consisted of two aspects: (1) whether the reported ability signal was sent to the opponent, and (2) whether individuals could opt out of the competition.

In the non-strategic baseline treatment, no ability signals were exchanged before the competition took place, and participants could not opt out of the competition. We compare this to the strategic treatment, where participants exchanged signals and could decide not to compete. In detail:

1. ***Non-Strategic Treatment (NStrat).*** In this treatment, participants were forced to compete and did not receive their opponents' ability signal before the competition. Thus, in each round, participants first stated their confidence[4] and then were entered into the competition and received feedback regarding whether they won or lost. Because of its non-strategic nature, we use this treatment as a baseline treatment against which we compare the other treatments.

2. ***Strategic Treatment (Strat).*** In each round of this treatment, a participant's reported confidence was shown to her randomly matched opponent as ability signal before the opponent decided whether to enter the tournament. This rule was known to all participants. After having observed the reported confidence of her opponent, each player chose whether to enter the competition, which consisted of the comparison of IQ scores, as explained above. If both players entered the competition (played In), the competitor with the higher score won 10 points, and the other player obtained zero

[4] As better explained later, the confidence level stated by players in (NStrat) were a truthful report, that we used to compute updating after wins and losses.

points.[5] Staying out of the competition resulted in a payoff of 5.5 points. A player who entered the competition and was paired with a player who opted out likewise received 10 points. Table 1 summarizes the payoff structure. Subsequently, each participant who entered the competition received feedback regarding whether she won or lost or whether the opponent did not enter. A player who did not enter did not learn whether the other player entered. A player who won uncontested also did not learn whether she was better or worse than the opponent. After each round, each player was re-matched with a randomly chosen opponent in her group of ten.

Table 1: Payoffs.

s_i \ s_j	*In*	*Out*
In	$I_i 10$, $(1 - I_i)10$	10, 0
Out	0, 10	5.5, 5.5

Notes: The table shows the payoffs in the Strategic treatment. I_i is an indicator variable that takes on the value 1 if row player i's score on the IQ test is higher than that of column player j and is zero otherwise. In the non-strategic treatment, opting out of the competition is not possible (i.e., the Out option is not relevant).

The possibility to signal ability to the competitor can create an incentive to overstate one's ability in the hope of deterring the opponent from entering altogether. Note, however, that bluffing by inflating signals beyond one's actual belief was costly in our setup. The total payoff from each round was the sum of the competition payoff plus the accuracy payoff based on the quadratic scoring rule. Bluffing meant decreasing the expected return from the perceived accuracy payoff.[6] We believe that this kind of signaling is an essential feature of real-world competitive settings. The same is true for the endogenous selection into competition. In many situations, one can choose whether to compete with someone. However, if competition does not take place, a definitive answer to the question "who is better?" is impossible.

These two treatments allow us to test the hypothesis that the strategic elements present in typical competitive environments (the presence of ability signals and endogenous entry

[5] In case of ties, a random device determined the winner.

[6] With the quadratic scoring rule, the cost of overconfidence is not linear in the deviation from true ability. Small deviations are only punished a little, but large deviations are punished considerably more. This reflects the fact that, whereas small exaggerations of one's ability may be somewhat socially acceptable, larger deviations may have more severe consequences when detected. This is similar to Kartik (2009), where lying costs stand for potential losses from being caught lying, such as reputation loss or social shame.

into competition) impact relative confidence. Clearly, however, they differ along two dimensions, each with a potential impact on player confidence. First, in the strategic treatment, participants can signal their own ability, which creates incentives for bluffing to deter the other player from entering the competition. The signals change the information set of the players across treatments. The received signals by themselves may affect the players' confidence update after winning or losing. For instance, losing (or winning) against someone who is perceived as very strong may have a different effect on one's own confidence than losing (or winning) against someone with similar stated confidence. Second, players can opt in and out of the competition and avoid receiving hard information about winning or losing. It is therefore difficult to say which of these elements drives the difference in outcome between our main two treatment variations.

To better understand the underlying mechanisms of confidence formation in a strategic context, we conducted two additional treatments, the *asymmetric treatments*. Although these additional treatments are not exhaustive in the sense that they cannot identify the importance of each possible mechanism, they do allow us to investigate specific mechanisms to gain insight into the underlying psychology of confidence formation due to (i) receiving opponent ability signals, (ii) a participant's own entry decisions, (iii) her opponent's entry decision. To this end, we created asymmetric treatments, in which we randomly assigned participants two different roles: Sender (S) and Receiver (R).

Senders could send an ability signal to Receivers, but could not opt out of the competition. Receivers could not send a signal to Senders but, after receiving a signal from their opponent, they could opt out of the competition. Otherwise, the structure and payoffs of the two asymmetric treatments follow the strategic treatment. In particular, the Sender would receive 10 points if the Receiver did not compete. Competition would only take place if the Receiver opted in, with all resulting payoffs as described in Table 1.

The difference between the two asymmetric treatments is the feedback generated within the game. In the first ***Asymmetric treatment without feedback (AsymNF)***, if R opted out, neither player learned who would have won the competition (as in the strategic treatment). By contrast, in the second treatment, ***Asymmetric treatment with Feedback (AsymF)***, player R learned whether she would have won or lost, even if she did not enter the competition, whereas player S got the same information as in the treatment without feedback.[7] Receivers in AsymF obtained the same feedback regarding the competitive outcomes as players in the baseline treatment. Yet, in contrast to the baseline players, player R also received the confidence signals of player S. The treatment can thus shed light on the

[7]From the perspective of player S, the two asymmetric treatments generate the same information, which is why we sometimes analyze the Senders of both treatments together in the following sections.

impact of receiving opponents' ability signals.

Stage 4: Final Confidence. In all experimental conditions, after the 12 rounds of Stage 3, we again elicited participant confidence (i.e., their beliefs about being in the top half of their group of ten players according to their IQ score). As before, correct beliefs were rewarded as per the quadratic scoring rule. Stage 4 had no strategic incentive to inflate elicited confidence.[8] This non-strategic measure thus allows us to identify how the different experimentally induced environments affect the relative overconfidence.

2.1 Data Collection and Procedure

We collected data from a total of 464 participants at the Laboratory for Behavioral Experiments at the University of Lausanne. Data were collected between 2013 and 2014. Participants were recruited from the regular subject pool by using ORSEE (Greiner, 2015) and mainly consisted of students from the University of Lausanne and from the Federal Polytechnic School of Lausanne.[9] The experimental software was programmed in z-Tree (Fischbacher, 2007). Treatments were randomized at the session level (see Table A1 in the Appendix for a randomization check comparing participant characteristics across treatments).

Upon entering the laboratory, participants drew a card from an opaque bag, assigning them their seat number. Participants received detailed written instructions (see Appendix B) before each stage of the experiment. While reading the instructions, participants could ask comprehension questions that experimenters answered in private. At the end of the experiment, participants filled in a post-experimental questionnaire containing demographics and some personality measures (see Table A1 in the Appendix).

The sessions lasted around 60 minutes. Participants received a show-up fee of 10 Swiss Francs. They earned additional money for the initial and final confidence beliefs in stages 2 and 4 (according to the quadratic scoring rule described above) and for a randomly selected round from stage 3 (belief and outcome of the tournament game). The exchange rate between points and Swiss Francs was 10 to 1. The average earnings were 28.20 Swiss Francs ($\sim$25 USD).

[8] Note that because the confidence measure elicited at this stage was not shown to any other participant, we exclude pure "presentation" signaling effects as documented by Proeger & Meub (2014) and Ewers & Zimmermann (2015), who show that people inflate their confidence when they know that their stated confidence will be seen by others.

[9] The subject pool includes undergraduates from all disciplines and we did not specify any exclusion restrictions for recruiting participants.

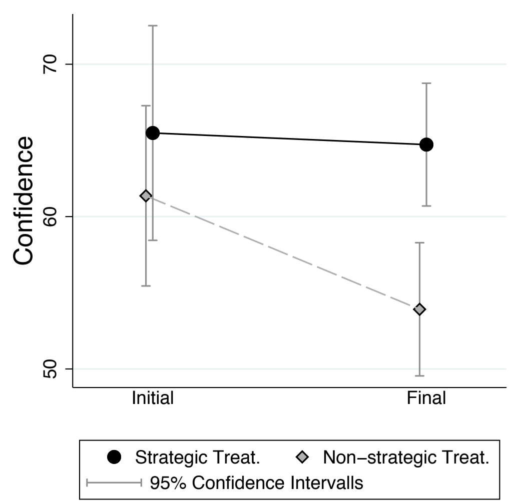
3 Results

We begin by analyzing the difference between the non-strategic and strategic treatments (Section 3.1) and show that overconfidence is sustained in the strategic treatment but not in the non-strategic treatment. We then turn to the asymmetric treatments to analyze the underlying mechanisms that lead to this difference (Section 3.2 and its subsections).

3.1 Main Effect of Treatment on Final Confidence

Figure 1 displays initial and final confidence in both the non-strategic and the strategic treatment. Before exposure to treatment, there is no difference in initial confidence [$t(174) = 1.11$, $p = 0.269$] and overconfidence prevails: Initial confidence is significantly above 50% in both the non-strategic (*NStrat*) [$t(95) = 4.60$, $p < 0.001$] and the strategic treatment (*Strat*) [$t(79) = 5.37$, $p < 0.001$].[10]

Figure 1: Initial and final confidence in the non-strategic and the strategic treatments.



This picture changes after exposure to treatments. In the non-strategic treatment, final confidence decreases significantly. Final confidence is 6.74 percentage points lower than initial confidence (signed rank test: $z = 2.24$, $p = 0.025$, $n = 8$),[11] and final confidence does not differ significantly from 50% (signed rank test: $z = 1.54$, $p = 0.124$, $n = 8$). This

[10] All p-values reported in the paper are for two-tailed tests.

[11] We always report inferential tests based on the smallest statistically independent unit. For final confidence this is the matching group, because of the interactions that took place between participants within a matching group during the experiment. Similarly, in regression analyses we cluster at the matching-group level whenever appropriate. For such regressions we also re-ran the analyses using wild bootstrap tests (Cameron & Miller, 2015), which confirmed the robustness of the results.

demonstrates that overconfidence is not sustained in a competitive environment where the possibility of strategic behavior is absent. In the strategic treatment, by contrast, overconfidence is sustained. Initial and final confidence levels are almost identical in the strategic treatment (signed rank test: $z = 0.14$, $p = 0.889$, $n = 8$) and final confidence remains significantly above 50% (signed rank test: $z = 2.52$, $p = 0.012$, $n = 8$). Comparing final confidence in the two treatment variations reveals a significant treatment difference (rank sum test: $z = 2.73$, $p = 0.006$, $n_1 = n_2 = 8$).

> **Main Result:** *Overconfidence is sustained in the strategic environment, whereas it decreases in the non-strategic environment, such that final confidence is significantly higher in the strategic treatment than in the non-strategic treatment.*

3.2 Why Is Final Confidence Higher in the Strategic Environment?

To this point, we have established that overconfidence is sustained in a strategic competitive environment in which people send ability signals and endogenously decide to enter, whereas confidence decreases after exposure to the non-strategic environment with forced competition and no signaling possibilities. What underlying mechanisms drive this treatment difference?

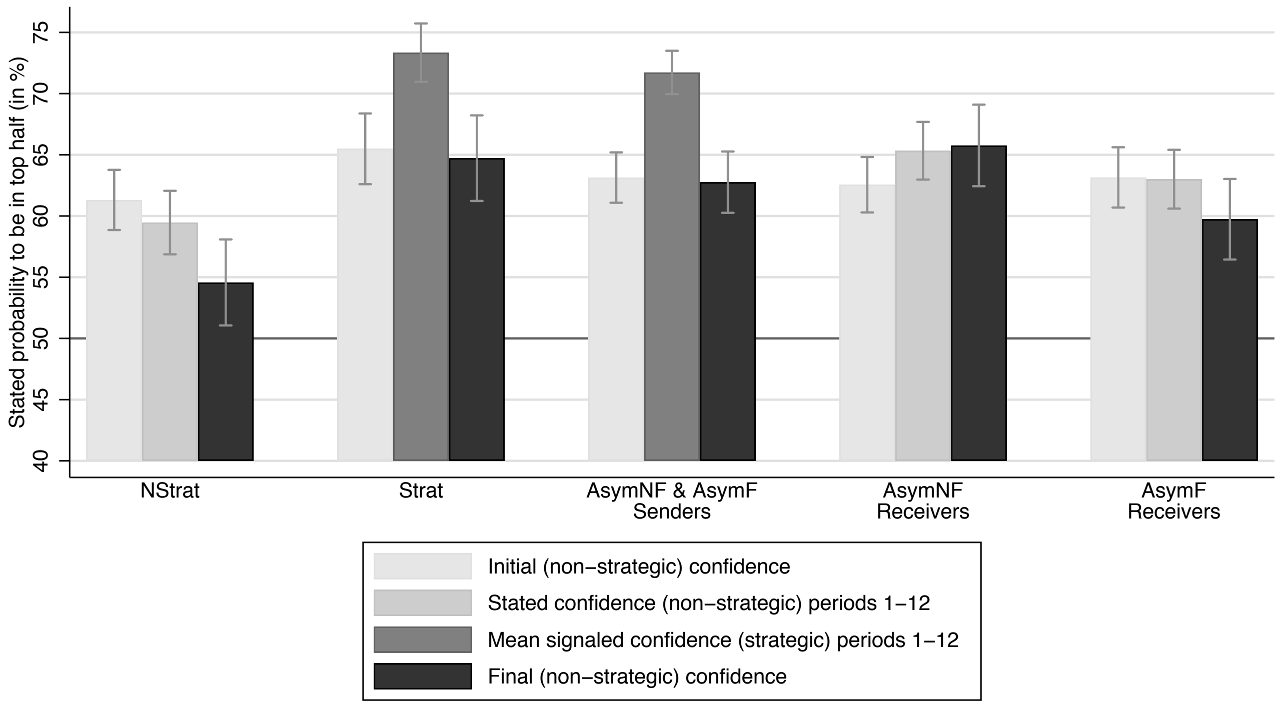
3.2.1 Bluffing and Deterrence

Figure 2 shows initial and final confidence by experimental condition. In addition, it also displays mean stated confidence across the 12 periods of the game in Stage 2. In the absence of a strategic incentive to inflate confidence, that is, when confidence cannot be used as a signal (as is the case in the baseline treatment and for Receivers in the asymmetric treatments), individuals report lower confidence on average in the 12 rounds of Stage 2 compared with Stage 1.

This contrasts with the strategic treatment and Senders in the asymmetric treatments. In these conditions, stated confidence serves as an ability signal: Individuals' stated confidence is shown to the opponents who can decide whether to enter competition. Figure 2 shows that, here, mean ability signals are higher than both initial and final confidence. On average, participants in the strategic treatment and Senders in the asymmetric treatments strategically inflate their ability signals by 8.33 percentage points (across periods 1–12) compared with initial confidence[12] and by 8.84 percentage points compared with final confidence. This

[12] Charness et al. (2018) obtained a similar result in a single-shot game.

Figure 2: Initial, signaled, and final confidence.



Note: Error bars represent plus/minus one standard error of the mean, clustered by individual.

demonstrates that individuals bluff by inflating ability signals, which is consistent with our theoretical model (in Appendix C).

Participants take the ability signals that they receive into account when deciding whether to enter into competition. In Table A2 of the Appendix, we regress entry decisions on a binary variable that indicates whether the received signal is higher than the participant's own confidence. If a participant's confidence is lower than the ability signal she receives, the probability of entering competition decreases by 36 percentage points (controlling for gender, IQ-score, and risk preferences). The regression also reveals that female participants put more weight on the received ability signals. For men, the probability of entering competition decreases by 25 percentage points when their confidence is lower than the received signal whereas, for women, it decreases by 45 percentage points (holding risk preferences and IQ score constant).

3.2.2 Confidence after Winning Unopposed: Senders

How do participants update their confidence when no competition occurred because their opponent stayed out? To investigate this mechanism we turn to Senders in the two asymmetric treatments. Other than in the strategic treatment, Senders in the asymmetric treatments do not receive ability signals and are forced to compete. However, they send ability signals and can experience that opponents are staying out of competition. This allows us to shed light on how unopposed winning affects final confidence.

Comparing initial and final confidence of the Senders in the asymmetric treatments (see *AsymNF* & *AsymF* Senders in Figure 2) shows that, on average, final confidence does not decrease compared with initial confidence (signed rank tests: *AsymNF* Senders: $z = 0.14$, $p = 0.674$, $n = 8$; *AsymF* Senders: $z = 0.14$, $p = 0.674$, $n = 8$). Initial and final confidence of Senders in the asymmetric treatments are virtually identical, mirroring the pattern found in the symmetric strategic treatment (*Strat*). Furthermore, the final confidence of Senders in the asymmetric treatments is significantly higher than the final confidence in the non-strategic baseline (rank sum tests: *AsymNF* Senders vs. *NStrat*: $z = 2.10$, $p = 0.036$, $n_1 = 8$, $n_2 = 8$; *AsymF* Senders vs. *NStrat*: $z = 1.79$, $p = 0.074$, $n_1 = 8$, $n_2 = 8$). Thus, the possibility of sending inflated signals and winning unopposed sustains confidence for Senders.

Table 2 shows regression results. In Column (1), we regress final confidence on treatment (and player-type) dummies. The non-strategic treatment is the omitted baseline treatment.[13] In Column (2), we analyze how final confidence depends on the events experienced by a participant during the experiment. We add the count measures "# Wins in Competition," which captures how many times a participant won in actual competition, and "# Wins Unopposed," which captures how many times a participant won because the opponent did not compete. These two variables allow us to investigate how updating differs between winning in competition and winning without an actual contest. For completeness, we control for "# Stay out," which captures how many times players in the strategic treatment and Receivers in the asymmetric treatment decided not to enter into competition. In Column (3), we add the variable "# Wins Unopposed X Would Lose," which captures how many times a player who would have lost in actual competition won only because the opponent stayed out. This event might be more likely for individuals who consciously inflate their ability signal to deter their opponents and hence are more cautious in updating their true confidence after winning because of an opponent staying out of competition. Including the variable "# Wins Unopposed X Would Lose" in the regression allows us to test for this possibility.

The regression results in Column (1) parallel our non-parametric tests. Final confidence is significantly higher for Senders in the asymmetric treatments compared with the non-

[13]To simplify the exposition, Table 2 contains only one indicator variable capturing sender types in both asymmetric treatments . Senders in both treatments were faced with almost identical instructions (the only difference in the instructions was that in the feedback treatment they were informed that the receiver types would learn who had the higher score in the IQ test even if a receiver decided not to enter). Importantly, we do not find any significant differences ($p > 0.10$) between the sender types in the two asymmetric treatments for initial and final confidence nor for signaled confidence in periods 1–12 (see also Table A3).

Table 2: Regression results for final confidence.

	DV: Final (Non-Strategic) Confidence		
	(1)	(2)	(3)
Strat	7.366***	6.466**	6.491**
	(2.500)	(3.093)	(3.092)
AsymNF / AsymF Senders	6.972**	−0.583	−0.824
	(3.011)	(3.782)	(3.885)
AsymNF Receivers	10.362***	14.415***	14.335***
	(2.221)	(3.062)	(3.050)
AsymF Receivers	3.935	9.234**	9.160**
	(2.914)	(3.714)	(3.696)
# Wins in Competition		5.096***	5.168***
		(0.441)	(0.431)
# Wins Unopposed		5.607***	5.490***
		(0.636)	(0.596)
# Wins Unopposed X Would Lose			0.752
			(1.136)
# Stay Out		0.558	0.597
		(0.568)	(0.557)
Initial Confidence (centered)	0.667***	0.336***	0.337***
	(0.055)	(0.063)	(0.064)
Constant	55.751***	24.324***	23.893***
	(1.866)	(3.089)	(3.046)
R^2	0.273	0.575	0.576
Observations	464	464	464

$^{*}p < .10,\ ^{**}p < .05,\ ^{***}p < .01$

Notes: OLS estimates. Final confidence (ranging from 0 to 100) is regressed on indicator variables for the different treatments or types. The non-strategic baseline treatment is the omitted treatment. "#" variables are counts of how many times the indicated event occurred during the 12 rounds of the experiment. "Initial Confidence" is centered at the sample mean. Robust standard errors clustered by 32 matching groups are in parentheses.

strategic baseline (see the coefficient 6.972 in Column (1)). Column (2) shows that the coefficients of both "# Wins in Competition" and "# Wins Unopposed" are virtually identical (post-estimation test: $\chi^2 = 1.90$, $p = 0.168$).[14] Similarly, Column (3) reveals that the coefficient of the interaction term "# Wins Unopposed X Would Lose" is not statistically significant. This is evidence that even those individuals who would have lost in actual competition do not account for this possibility and update just like winning in actual competition. Thus, the difference in final confidence between the non-strategic treatment and the Senders in the asymmetric treatments is driven by Senders who would have lost in actual competition but who won because their opponents stayed out (i.e., the successful bluffers).[15] Columns (2) and (3) also reveal that controlling for Senders' experiences during the experiment fully accounts for the treatment differences. The dummy variable "AsymNF / AsymF Senders," which captures the remaining treatment differences, is not statistically different from zero.[16]

> **Result 1:** *Participants, even those who would have lost in actual competition, update their confidence in response to an unopposed win as if they had won in actual competition.*

3.2.3 Inflated Signals, Staying Out, and Confidence: Receivers

Here we investigate (i) how individuals update their confidence after they stayed out of competition and (ii) whether received ability signals directly impact individuals' confidence. Focusing on Receivers in the asymmetric treatments allows us to address these questions because Receivers obtain ability signals and endogenously decide to enter competition.

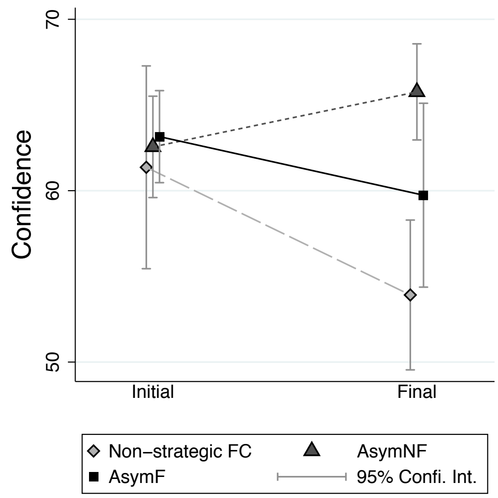
Comparing Receivers' initial and final confidence in the asymmetric treatment without feedback (*AsymNF*) in Figure 3 shows that their confidence does not decrease during the

[14] Senders in *AsymNF / AsymF* update very similarly after winning in competition as participants in *NStrat*. Entering additional interaction terms of "# Wins in Competition" with treatment dummies yields non-significant coefficients for "# Wins in Competition X AsymNF / Asym F Senders" ($p = 0.770$) and for "# Wins in Competition X AsymNF Receivers" ($p = 0.177$). The interaction "# Wins in competition X AsymF Receivers" is marginally significant ($p = 0.059$) in the direction that *AsymF* Receivers update somewhat less strongly after winning in competition than do participants in *NStrat*.

[15] In the asymmetric treatments, Senders win more often than in *NStrat* because, contrary to the prediction of an equilibrium model (see Appendix C), Receivers are not always capable to correctly infer whether they would lose or win by entering the contest. In the baseline *NStrat* treatment, players receive on average six times the feedback that they win. The Senders in the *AsymNF / AsymF* treatments win on average three times opposed and on average 4.23 times unopposed (in total 7.23). Given that players in the *NStrat* treatment update after winning as Senders in the asymmetric treatments, the additional effect ($1.23 \times 5.7 = 7.011$) on final confidence in "AsymNF / AsymF" explains the entire difference between the baseline and the Senders in AsymNF / AsymF [the coefficient 6.972 in Column (1)].

[16] More specifically, this suggests that sending inflated signals *per se* does not have a direct effect on the sender's own confidence; a factor we cannot rule out in the non-parametric tests since two dimensions change between the *non-strategic* treatment and Senders in the asymmetric treatments; namely, (1) that the latter can send signals, and (2) that the opponents may stay out.

Figure 3: Initial and final confidence in the non-strategic baseline and Receivers in the two asymmetric treatments.



experiment. Rather, it increases insignificantly by 3.21 percentage points (signed rank test: $z = 1.54$, $p = 0.124$, $n = 8$). Consequently, final confidence is significantly greater than in the baseline treatment (rank sum test: $z = 3.37$, $p = 0.001$, $n_1 = n_2 = 8$) and overconfidence prevails as final confidence remains significantly above 50% (signed rank test: $z = 2.53$, $p = 0.012$, $n = 8$).

The difference between the baseline treatment and the Receivers in the asymmetric treatment without feedback may be driven by at least two effects: (i) receiving less feedback due to opting out of competition and (ii) being exposed to the inflated ability signals of their opponents. Our asymmetric treatment with feedback (*AsymF*) allows us to disentangle these two channels by eliminating the first possibility. In this treatment, we informed Receivers who opted out of competition whether they would have won or lost. With regard to the information set, the only difference compared with the baseline treatment is thus that Receivers in *AsymF* were exposed to the potentially inflated ability signals sent by their opponents.

Figure 3 shows that final confidence in the *AsymF* treatment falls between confidence in the non-strategic baseline treatment and in the asymmetric treatment without feedback (*AsymNF*). Non-parametric tests reveal that the decrease between initial and final confidence in the *AsymF* treatment is not significant (signed rank test: $z = 0.98$, $p = 0.327$, $n = 8$). However, the receivers' final confidence in *AsymF* is significantly lower than in *AsymNF* (rank sum test: $z = 2.10$, $p = 0.035$, $n_1 = n_2 = 8$), whereas it is significantly greater than in the baseline and also significantly above 50% (signed rank test: $z = 2.52$, $p = 0.012$, $n = 8$). This suggests that feedback to Receivers in *AsymF* does matter, but it cannot explain the full treatment difference.

Table 3: Regression results for round-by-round confidence updating (*Non-strategic* baseline and Receivers in *AsymNF* and *AsymF* only).

	DV: Change in confidence			
	(1)	(2)	(3)	(4)
Won	1.714***	1.714***	1.891***	1.706***
	(0.288)	(0.288)	(0.313)	(0.295)
Stayed out	1.136***	1.125*	0.650	0.658
	(0.374)	(0.575)	(0.538)	(0.671)
Lost	−3.711***	−3.711***	−3.651***	−3.548***
	(0.528)	(0.528)	(0.507)	(0.491)
Stayed Out X Feed/b Won		2.025	1.929	2.142
		(1.792)	(1.713)	(1.384)
Stayed Out X Feed/b Lost		−0.819	−0.753	−1.519*
		(0.920)	(0.849)	(0.816)
Oppnt's Sign Confidence (cent.)			0.048***	0.131***
			(0.013)	(0.037)
Oppnt's Sign Conf X Won				−0.134***
				(0.039)
Oppnt's Sign Conf X Stayed Out				−0.084*
				(0.048)
Oppnt's Sign Conf X Stayed Out X Feed/b Won				−0.018
				(0.052)
Oppnt's Sign Conf X Stayed Out X Feed/b Lost				0.090
				(0.058)
R^2	0.024	0.025	0.028	0.033
Observations	2880	2880	2880	2880

$^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$

Notes: OLS estimates. The change in confidence from one round to the next is the dependent variable (DV). Only observations from the *non-strategic* baseline treatment and from Receivers in the *AsymNF* and *AsymF* treatments are included. The constant is suppressed. "Won," "Stayed Out," and "Lost" indicate whether a participant won, stayed out, or lost in the round preceding the confidence update. In Column (2), we include indicator variables of whether a participant, who stayed out in the *AsymF* treatment, received feedback to have lost or to have won. In Column (3), we add the opponent's signal received by a participant in the *AsymNF* and *AsymF* treatments ("oppnt's" is short for "opponent's"). In Column (4), we interact the opponent's signal received with a participant's experience prior to his confidence update [i.e., with all variables of Column (2)]. Standard errors clustered by 24 matching groups are reported in parentheses.

In the regression analyses reported in Table 3 we provide evidence that inflated signals have a direct effect on confidence. We regress the change in confidence ($c_{t+1} - c_t$) between any two of the 12 rounds of the experiment on their experience in a given round t. Other than in the previous subsection, we can use information on confidence for each round because we only included data of the baseline and of Receivers in the two asymmetric treatments in the regression analysis. In these treatments, there are no strategic incentives for inflating confidence during the experiment, and truth-telling is incentivized via the quadratic scoring rule. We thus get a non-inflated confidence measure for each round.[17]

In Column (1), we regress the change in confidence on experiencing winning, experiencing losing, or staying out. Note that we suppressed the constant in the models. The coefficients for "Won," "Stayed out," and "Lost" thus reflect the update in response to each of those events. In Column (2), we use the data from the asymmetric treatment with feedback (*AsymF*) and include two interactions: "Stay out X Feedback Won" and "Stay out X Feedback Lost." In Column (3), we include "Oppnt's Signaled Confidence" (short for "Opponent's Signaled Confidence") which is the mean-centered ability signal that the Receiver got from her opponent. This allows us to analyze the impact of the message received on player confidence. Since this impact may vary with past experiences in the game, in Column (4), we have this signal interact with the explanatory variables from Column (2). This specification allows us to investigate whether participants account for the message differently depending on the outcome of the previous competition.

The confidence of those who "Stayed out" does not decrease, as can be seen in all four columns of Table 3.[18] Thus, in contrast to Senders, whose confidence increases after an uncontested win, there is no associated decrease for Receivers, who stay out of competition.[19]

> **Result 2:** *Participants who stay out of competition do not update their confidence downwards.*

[17]Note that, for the last of the 12 rounds, we used the final confidence measure of Stage 4 as c_{t+1}. This gives us 12 observations for each participant.

[18]In fact, Columns (1) and (2) show even significantly positive coefficients for "Stayed out." Columns (3) and (4) offer an explanation for this finding: those who stay out increase their confidence based on the ability signals they receive from their opponents. Once we control for the received signals in Columns (3) and (4) the coefficients for "Stayed out" are no longer significant.

[19]To account for individuals who are already at the upper bound of 100 percent or at the lower bound of 0 percent and have constraints on the scope of updating, we refer to Table A4 in the Appendix. There, we only include observations where increasing or decreasing confidence is still feasible (i.e., where $c_t < 100$ or $c_t > 0$). Our main takeaway from this regression stands. People do not update downwards after staying out of competition. At the same time, Table A4 reveals that not too much should be read into the coefficients for "Won" and "Lost." Excluding extreme-value observations shows that updating upwards after winning and updating downwards after losing are symmetric in this case (i.e., the extent of updating downwards or upwards is almost identical).

The opponent's signaled confidence is, actually, significantly positively associated with increased confidence, as shown in Column (3). This is evidence that Receivers take the inflated ability signals of their opponents into account when updating their own ability. Column (4) provides further insights. The effect of the opponent's signaled confidence is strongest and clearly significant when a Receiver lost in competition in round t (captured by the main effect of "Opponent's Signaled Confidence (cent)" in Column (4)). The marginally significant coefficient for the interaction term "Oppnt's Sign Conf X Stayed Out" indicates that this effect is less strong when staying out, and a test of joint significance indicates that it is indeed not statistically significant in this case (post-estimation test for joint significance of "Oppnt's Signaled Confidence (cent)" and "Oppnt's Signaled Conf X Stayed out": $F(1, 23) = 2.53$, $p = 0.125$). Similarly, in case of winning, the opponent's confidence signal also does not have a significant influence on player confidence (post-estimation test for joint significance of "Oppnt's Signaled Confidence (cent)" and "Oppnt's Signaled Conf X Won" $F(1, 23) = 0.05$, $p = 0.831$). These findings do not change upon applying the alternative specifications reported in Table A4 in the Appendix. Thus, if a participant lost in competition, the confidence update positively depends on the opponent's signaled confidence. This points to the third reason for sustained overconfidence in our setup:

> **Result 3:** *Inflated confidence signals are "contagious," because the higher the opponent's signal, the higher the confidence update of participants who lose in competition.*

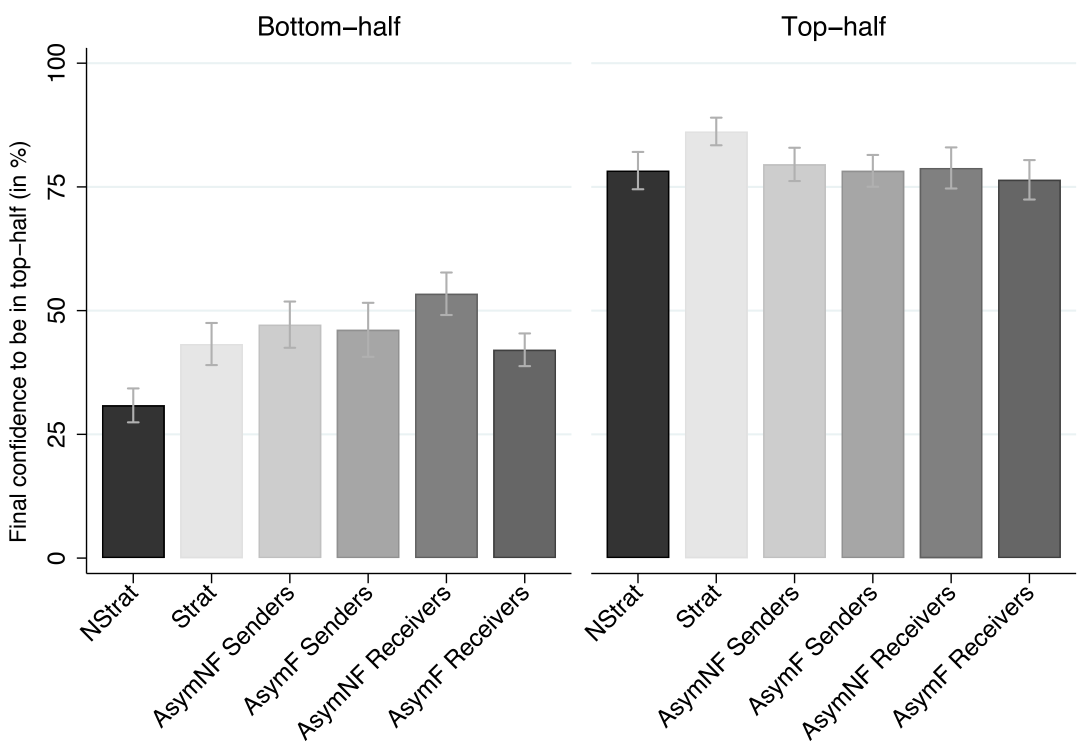
4 Heterogeneity

Our findings suggest that accumulated experiences in past strategic interactions shape individuals' confidence. These experiences depend on individual behavior in the strategic environment (i.e., they are closely linked to an individual's propensity to compete and/or to deter opponents). Behavioral differences within the strategic environment can thus lead to heterogeneity in confidence. In the next two sections, we investigate whether exposure to strategic competitive environments can shape heterogeneity in confidence in a way that is consistent with two well-established empirical regularities: Low-skilled individuals are more overconfident than high-skilled individuals (the Dunning–Kruger effect), and women are less confident than men.

4.1 Dunning–Kruger

Kruger & Dunning (1999) observed that overconfidence is more common among low-skilled individuals. They attribute this finding to a lack of metacognitive skills (i.e., that low-skilled individuals lack the cognitive ability to realize their limitations). Our study provides evidence that overconfidence of those with low skills can also arise from differential learning about one's own skill in a competitive strategic environment. At the same time, we find no evidence that metacognitive skills vary with regard to learning about one's relative standing.

Figure 4: Final confidence for bottom-half and top-half participants.



Note: Error bars represent plus or minus one standard error of the mean, clustered by individual participants.

Figure 4 displays final confidence by treatment for individuals in the bottom half of their reference group of ten (left panel) and in the top half (right panel). The figure reveals that the treatment differences stem almost entirely from the bottom-half participants, which could be interpreted as evidence for the Dunning–Kruger effect. However, we argue that this would be a misinterpretation and that, instead, the mechanisms discussed in the previous section have a stronger effect on low-performing individuals.

First, we have shown that Senders update their confidence in a similar way after winning in actual competition and after winning unopposed. For a player who would have won anyway (more likely a high-skilled player), this does not distort their confidence. However, it fosters overconfidence for Senders who would have lost in actual competition but were paired with an opponent who did not compete. This is more often the case for low-skilled Senders. Indeed, Figure 4 reveals that final confidence for low-skilled Senders in the asymmetric treatments is considerably higher than in the non-strategic baseline treatment. Second, individuals who stay out of the competition do not update their confidence. This has a differential effect by skill. Bottom-half players who stay out of the competition have a higher propensity to lose in actual competition than top-half players. By staying out, they do not learn that they would have lost in actual competition and therefore do not update their ability downwards. Third, individuals who have lost or who stayed out rely more strongly on the inflated signals they receive when updating their confidence, as demonstrated in the previous section. Losing and staying out is more common among low-skilled players, which further contributes to sustained overconfidence among the low-skilled players. The latter two points are consistent with Figure 4, where the confidence of low-skilled Receivers does not decrease in the asymmetric treatment without feedback and only does so partially when feedback is provided. Thus, all three channels summarized by Results 1–3 foster confidence for low-skilled individuals. In contrast, the decrease in final confidence in the non-strategic treatment, where these channels are absent, is almost fully driven by the low-skilled players. This constitutes evidence that low-skilled players can update their relative-ability estimate and thus do not lack the metacognitive skills to realize their limitations. Instead, it is the strategic setting that leads to their persistence of overconfidence.

In Table A5 of the Appendix, we provide regression results in which we analyze confidence updating of top- and bottom-half participants in more detail. The regression results corroborate those of Figure 4, and we provide further evidence that bottom- and top-half individuals have similar metacognitive skills for updating their relative ability. Ability and confidence determine the frequency with which an individual experiences the different events of winning, losing, and not competing. Yet, conditional on these events, updating is very similar among top- and bottom-half participants.

4.2 Gender

A large body of literature provides evidence that men are more overconfident than women. The results of the present study are consistent with this finding. Initial confidence to be in the top half for men is 71 percent, whereas for women it is only 55 percent (rank sum: $z = 8.016, p < 0.001, n = 464$). Our study emphasizes that experiences in strategic competitive environments shape people's confidence. Can gender differences in confidence arise from different experiences in strategic competitive environments?

First, we find that men deter their opponents more often in cases where they would have lost in an actual competition than women. Empirical studies suggest that men are generally more confident but also more dishonest than women (see, e.g., Dreber & Johannesson, 2008). In the strategic treatment and for Senders in the two asymmetric treatments, men who would have lost in actual competition win unopposed in about 10% of the cases. For women, this event occurs only in about 7% percent of the cases. Although statistically significant ($\chi^2 = 7.28$, $p = 0.007$), this difference is quantitatively rather small.[20]

Second, when given the option, women opt out of competition 40% of the time compared to 25% of the time for men. As a consequence of the endogenous selection, women experience winning in competition 36% of the time (47% for men; rank sum: $z = 3.489, p < 0.001, n = 48$), while they lose 25% of the time (27% for men; rank sum: $z = 0.413, p = 0.680, n = 48$). There are no significant gender differences in the experience of losing, whereas a smaller fraction of women than men experience winning when they choose to enter into a competition. Were women forced to compete, they would experience winning about 28 percent more often compared to what they do in the endogenous selection treatments. Men would experience winning only 9 percent more often. This finding is consistent with results in the literature that has established that women tend to avoid competition more than men (see, e.g., Gneezy et al., 2003; Gneezy & Rustichini, 2004; Niederle & Vesterlund, 2007; Croson & Gneezy, 2009). We hypothesize that not competing has a differential impact on women's confidence relative to men depending on ability. For high-skilled women, avoidance of competition will tend to decrease confidence because they experience winning less often compared with high-skilled men, who enter more frequently. Conversely, low-skilled women will experience losing less often compared with men, which tends to increase confidence. And third, we find that women do not condition their confidence to a larger degree on the received signal from the opponent than men (see Table A6 in the Appendix).

In the regression analysis in Table A7 in the Appendix, the dependent variable is the final non-strategic confidence. We find that a female Sender in the top half of her group is significantly less confident than her male counterpart (see coefficient on "AsymSender X Fem X Top" in Columns (2) and (3)). However, we find no other conclusive evidence in Table A7: the coefficients for the gender-interaction terms go in the expected direction, but they are not significant.

One possible explanation is that women's initial average assessment of being in the top

[20]Collapsing at the matching-group level and testing with a rank sum test still reveals a marginally significant difference (rank sum: $z = 1.953, p = 0.051, n = 48$).

half is 55 percent, which is not significantly different from 50 percent, which is expected in a population without overconfidence. Because our strategic treatments make initial overconfidence persistent, without being able to induce people to improve their assessment, we could conclude that, since women are accurate in their initial assessment, there is a smaller range for adjustment.

5 Discussion and Concluding Remarks

In this article, we tested the hypothesis that strategic environments foster overconfidence. We provided laboratory evidence that, in a competitive strategic environment, overconfidence persists, whereas in a non-strategic competitive environment, overconfidence disappears.

We investigated several mechanisms that may explain this result. Consistent with our theoretical model, we find that, in a strategic environment, ability signals are inflated. At the same time, players take these inflated signals into account when deciding whether to opt in or out of competition. This behavior in the strategic environment is compatible with bluffing and deterrence and changes the accumulated feedback that individuals receive, which ultimately impacts individual confidence. However, the difference between the strategic and non-strategic treatments is not merely due to the fact that opting out of competition reduces the amount of information individuals receive in the strategic environment. We highlight three mechanisms: First, individuals who would have lost in actual competition but solely won because their opponent stayed out update just as if they won in actual competition. Second, individuals who do not compete do not update their confidence. Third, the inflated signals are "contagious." Individual confidence increases due to the inflated ability signals that they receive.

These mechanisms provide evidence that heterogeneity in confidence may originate from differences in accumulated feedback that individuals received in past strategic interactions. Apart from their actual ability, this feedback depends on their behavior in the strategic environment, which is closely linked to an individual's propensity to enter competitions and/or to deter opponents. For example, a person who inflates her ability is more likely to successfully deter others. Successful deterrence results in higher confidence, which may further enhance the credibility of future ability signals (Schwardmann & van der Weele, 2019; Trivers, 2011).

These channels can add a different angle to the empirical finding of the Dunning–Kruger effect. Rather than lacking the metacognitive skills to update their relative standing, low-skilled individuals have different experiences in a strategic environment and hence accumulate different feedback about their relative standing. Low-skilled individuals are more likely to opt out of competition and thus receive less feedback that they would have lost against an opponent. At the same time, low-skilled individuals who are good at deterring their opponents likewise do not get information that they would have lost in actual competition. This is not related to differences in metacognitive skills. Conditional on these experiences, there is no difference in updating by skill. The present work provides empirical support for this interpretation. We show that the differences in treatment are mainly driven by low-skilled individuals.

We also investigated whether these mechanisms can explain gender differences in overconfidence. Although no conclusive evidence exists overall, evidence does exist that women's behavior in a competitive environment is less conducive to foster overconfidence compared with that of men. First, we provide evidence that women are less likely to win by bluffing (i.e., there are fewer instances where women who would have lost in actual competition win because their opponent did not compete). Second, compared with men, a higher fraction of women do not experience winning because they opted out of competition.

The focus on accumulated feedback highlights the role of starting conditions. Our findings are therefore consistent with those of Belmi et al. (2019), who propose that social class shapes the attitudes that people hold about their abilities. They show that individuals with relatively high social class are more overconfident compared with individuals with relatively low social class. Although the present study does not provide direct evidence to support this claim, it is plausible that members of higher social classes are challenged less often in competition because they are shielded by their high class affiliation. Consequently, "high-class" individuals do not learn that they would lose in actual competition, whereas those who are deterred by high class affiliation do not learn that they would have won in actual competition.

Similarly, the mechanisms outlined here may explain why overconfidence is prevalent among leaders. Leaders may be particularly able to display signals of high ability. As a consequence, they are less often challenged in actual competition and may mistakenly attribute the lack of challengers to their ability, not accounting for the possibility that they would have lost had competition actually occurred.

References

Banerjee, S., Humphery-Jenner, M., Nanda, V., & Tham, M. (2018). Executive overconfidence and securities class actions. *Journal of Financial and Quantitative Analysis*, *53*(6), 2685–2719.

Barber, B. & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*, *116*(1), 261–292.

Belmi, P., Neale, M. A., Reiff, D., & Ulfe, R. (2019). The social advantage of miscalibrated individuals: The relationship between social class and overconfidence and its implications for class-based inequality. *Journal of personality and social psychology*.

Bénabou, R. & Tirole, J. (2002). Self-confidence and personal motivation. *Quarterly Journal of Economics*, *117*(3), 871–915.

Benoît, J. & Dubra, J. (2011). Apparent overconfidence. *Econometrica*, *79*(5), 1591–1625.

Benoît, J.-P., Dubra, J., & Moore, D. (2015). Does the better-than-average effect show that people are overconfident?: Two experiments. *Journal of the European Economic Association*, *13*(2).

Bertrand, M. (2011). New Perspectives on Gender. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics*, volume 4B chapter 17, (pp. 1543–1590).

Brunnermeier, M. K. & Parker, J. A. (2005). Optimal expectations. *American Economic Review*, *95*(4), 1092–1118.

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

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

Cameron, C. A. & Miller, D. L. (2015). A practitioner?s guide to cluster-robust inference. *Journal of Human Resources*, *50*(2), 317–372.

Charness, G., Rustichini, A., & Van de Ven, J. (2018). Self-confidence and strategic behavior. *Experimental Economics*, *21*(1), 72–98.

Compte, O. & Postlewaite, A. (2004). Confidence-enhanced performance. *American Economic Review*, 1536–1557.

Croson, R. & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, *47*(2), 1–27.

Dreber, A. & Johannesson, M. (2008). Gender differences in deception. *Economics Letters*, *99*(1), 197–199.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology*, *57*(6), 1082–1090.

Ewers, M. & Zimmermann, F. (2015). Image and misreporting. *Journal of the European Economic Association*, *13*(2), 363–380.

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

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

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

Gneezy, U. & Rustichini, A. (2004). Gender and competition at a young age. *American Economic Review*, *94*(2), 377–381.

Goel, A. M. & Thakor, A. V. (2008). Overconfidence, CEO selection, and corporate governance. *Journal of Finance*, *63*(6), 2737–2784.

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

Hambrick, D. C. (2007). Upper echelons theory: An update. *Academy of Management Revie*, *32*(2), 334–343.

Hambrick, D. C. & Mason, P. A. (1984). Upper echelons: The organization as a reflection of its top managers. *Academy of management review*, *9*(2), 193–206.

Hiller, N. J. & Hambrick, D. C. (2005). Conceptualizing executive hubris: the role of (hyper-) core self-evaluations in strategic decision-making. *Strategic Management Journal*, *26*(4), 297–319.

Hirshleifer, D., Low, A., & Theo, S. H. (2012). Are overconfident ceos better innovators? *Management Science*, *67*(4), 1457–1498.

Johnson, D. D. (2009). *Overconfidence and war: The havoc and glory of positive illusions*. Harvard University Press.

Kartik, N. (2009). Strategic communication with lying costs. *The Review of Economic Studies*, *76*(4), 1359–1395.

Kőszegi, B. (2006). Ego utility, overconfidence, and task choice. *Journal of the European Economic Association*, *4*(4), 673–707.

Kruger, J. & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121.

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

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

Merkle, C. & 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.

Moebius, M. M., Niederle, M., Niehaus, P., & Rosenblat, T. S. (2014). Managing self-confidence. *Working Paper*.

Muthukrishna, M., Henrich, J., Toyakawa, W., Hamamura, T., Kameda, T., & Heine, S. J. (2018). Overconfidence is universal? elicitation of genuine overconfidence (ego) procedure reveals systematic differences across domain, task knowledge, and incentives in four populations. *PLoS ONE*, *13*(8), e0202288.

Nadkarni, S. & Herrmann, P. (2010). Ceo personality, strategic flexibility, and firm performance: The case of the indian business process outsourcing industry. *Academy of Management Journal*, *53*(5), 1050–1073.

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

Peterson, R. S., Smith, D. B., Martorana, P. V., & Owens, P. D. (2003). The impact of chief executive officer personality on top management team dynamics: one mechanism by which leadership affects organizational performance. *Journal of Applied Psychology*, *88*(5), 795.

Phua, K., Tham, T. M., & Chishen, W. (2018). Are overconfident ceos better leaders? evidence from stakeholder commitments. *Journal of Financial Economics*, *127*(3), 519–545.

Proeger, T. & Meub, L. (2014). Overconfidence as a social bias: Experimental evidence. *Economics Letters*, *122*(2), 203–207.

Raven, J. (2000). *Raven's Advanced Progressive Matrices*. Pearson.

Scheinkman, J. & Xiong, W. (2003). Overconfidence and speculative bubbles. *Journal of Political Economy*, *111*(6), 1183–1219.

Schwardmann, P. & van der Weele, J. (2019). Deception and self-deception. *Nature Human Behavior*, 1–7.

Trivers, R. (2011). *Deceit and Self-Deception: Fooling Yourself the Better to Fool Others*. Penguin.

Wang, G., Holmes Jr, R. M., Oh, I.-S., & Zhu, W. (2016). Do ceos matter to firm strategic actions and firm performance? a meta-analytic investigation based on upper echelons theory. *Personnel Psychology*, *69*(4), 775–862.

Weinberg, B. (2009). A model of overconfidence. *Pacific Economic Review*, *14*(4), 502–515.

Zimmermann, F. (2019). The dynamics of motivated beliefs. *American Economic R.*

Appendix

A Additional Results and Analyses

Table A1: Randomization tests.

	NStrat	Strat	AsymNF & F Senders	AsymNF Receivers	AsymF Receivers
IQ Quiz Score	8.271	8.238	8.194	8.403	8.125
	(2.494)	(2.430)	(2.703)	(2.866)	(2.534)
Initial Confidence	61.313	65.487	63.139	62.556	63.153
	(24.074)	(25.818)	(24.698)	(19.210)	(20.903)
Age	21.052	21.262	21.090	21.583	20.875
	(2.841)	(2.438)	(2.826)	(2.949)	(2.926)
Female	0.531	0.475	0.458	0.542	0.556
	(0.502)	(0.503)	(0.500)	(0.502)	(0.500)
Foreign student	$0.490_{a,b}$	0.388_{a}	$0.549_{a,b}$	0.389_{a}	$0.486_{a,b}$
	(0.503)	(0.490)	(0.499)	(0.491)	(0.503)
Bachelor	0.823_{a}	0.850_{a}	$0.764_{a,b}$	$0.750_{a,b}$	$0.889_{a,c}$
	(0.384)	(0.359)	(0.426)	(0.436)	(0.316)
Master	0.135_{a}	0.113_{a}	0.132_{a}	$0.208_{a,b}$	$0.083_{a,c}$
	(0.344)	(0.318)	(0.340)	(0.409)	(0.278)
Other than Bachelor or Master	0.042_{a}	0.037_{a}	0.104_{b}	$0.042_{a,b}$	0.028_{a}
	(0.201)	(0.191)	(0.307)	(0.201)	(0.165)
Business, Econ & Law	0.385	0.275	0.403	0.389	0.347
	(0.489)	(0.449)	(0.492)	(0.491)	(0.479)
Social Science & Humanities	0.094_{a}	0.263_{b}	$0.146_{a,c}$	$0.153_{a,b,c}$	$0.222_{a,b,c}$
	(0.293)	(0.443)	(0.354)	(0.362)	(0.419)
Medicine & Biology	0.125_{a}	0.113_{a}	$0.076_{a,b}$	0.014_{b}	$0.056_{a,b}$
	(0.332)	(0.318)	(0.267)	(0.118)	(0.231)
Environmental Science & Architecture	0.156	0.075	0.139	0.097	0.167
	(0.365)	(0.265)	(0.347)	(0.298)	(0.375)
Engineering & IT	0.146	0.175	0.167	0.222	0.111
	(0.355)	(0.382)	(0.374)	(0.419)	(0.316)
Other study subject	0.094	0.100	0.069	0.125	0.097
	(0.293)	(0.302)	(0.255)	(0.333)	(0.298)
# participants known previously	$0.573_{a,b}$	$0.588_{a,b}$	$0.611_{a,b}$	0.333_{a}	0.819_{b}
	(0.926)	(0.964)	(0.854)	(0.628)	(2.480)
Familiar w/ cond. probability (0/1)	0.490	0.412	0.528	0.431	0.389
	(0.503)	(0.495)	(0.501)	(0.499)	(0.491)
Plays competitive sports (0/1)	0.271	0.263	0.292	0.333	0.292
	(0.447)	(0.443)	(0.456)	(0.475)	(0.458)
Agreeableness (1-5)	3.771	3.925	3.899	3.878	3.816
	(0.694)	(0.704)	(0.572)	(0.649)	(0.586)
Extraversion (1-5)	3.435	3.403	3.457	3.385	3.552
	(0.722)	(0.855)	(0.692)	(0.769)	(0.816)
Neuroticism (1-5)	2.870	2.794	2.736	2.719	2.878
	(0.864)	(0.723)	(0.835)	(0.884)	(0.808)
Openness (1-5)	3.628	3.809	3.620	3.628	3.622
	(0.759)	(0.720)	(0.727)	(0.838)	(0.784)
Conscientiousness (1-5)	3.339	3.309	3.217	3.198	3.188
	(0.973)	(0.891)	(0.962)	(0.890)	(0.935)
Honesty (1-5)	3.401	3.444	3.404	3.417	3.414
	(0.658)	(0.530)	(0.598)	(0.566)	(0.610)
Core Self-Evaluations (1-5)	3.372	3.283	3.394	3.394	3.306
	(0.612)	(0.619)	(0.583)	(0.620)	(0.536)
Risk Propensity (0-17)	6.260	6.938	6.403	6.569	6.694
	(2.967)	(3.985)	(3.167)	(3.094)	(3.837)
Cognitive Reflection Test (0-3)	1.333_{a}	1.337_{a}	$1.167_{a,b}$	$1.194_{a,b}$	0.986_{b}
	(1.073)	(1.158)	(1.044)	(1.083)	(1.028)
Social Desirability (0-25)	15.125_{a}	$16.063_{a,b}$	16.424_{b}	$14.958_{a,c}$	$16.028_{a,b,c}$
	(3.319)	(3.750)	(3.087)	(3.083)	(3.435)
N (participants)	96	80	144	72	72

Notes: The table reports means. Standard deviations (calculated at the participant level) are in parentheses. Values within a row that do not share a subscript differ significantly ($p < 0.05$). In rows without subscripts, there are no significant differences for all possible comparisons within the row. The overall F-test is significant ($p < 0.05$) only for "Social Science & Humanities." Senders in the *AsymNF* and *AsymF* treatments were collapsed in accordance with the analyses presented in the paper (see footnote 13).

Table A2: Regression results for entry decisions.

	DV: Entering Competition		
	(1)	(2)	(3)
Higher Signal	-0.429^{***}	-0.358^{***}	-0.253^{***}
	(0.023)	(0.027)	(0.042)
Female		-0.040	0.114
		(0.040)	(0.108)
HigherSignal x Female			-0.192^{***}
			(0.064)
IQ Score		0.042^{***}	0.041^{***}
		(0.006)	(0.007)
IQ Score x Female			0.003
			(0.008)
Risk preference		0.007^{**}	0.013^{***}
		(0.003)	(0.004)
Risk Preference x Female			-0.011
			(0.008)
Constant	0.889^{***}	0.476^{***}	0.389^{***}
	(0.018)	(0.060)	(0.084)
R^2	0.199	0.261	0.272
Observations	1728	1728	1728

$^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$

Notes: Linear probability model. Entry decision is regressed on the binary variable that indicates whether the received opponent's signal is higher than one's own confidence. Included are all decisions of Receivers in the two asymmetric treatments. Additional covariates in Column (2) are gender, risk preferences, and ability. In Column (3), all covariates are made to interact with gender. Robust standard errors clustered at matching groups are reported in parentheses.

Table A3: Descriptive statistics.

	NStrat	Strat	AsymNF Senders	AsymF Senders	AsymNF Receivers	AsymF Receivers
All participants:						
IQ Quiz Score	8.27	8.24	7.96	8.43	8.40	8.13
	(2.49)	(2.43)	(2.94)	(2.44)	(2.87)	(2.53)
Initial Confidence	61.31	65.49	61.69	64.58	62.56	63.15
	(24.07)	(25.82)	(24.31)	(25.17)	(19.21)	(20.90)
Periods 1-12: Mean Confidence	59.46	73.34	71.09	72.37	65.33	63.01
	(25.42)	(21.33)	(21.14)	(21.43)	(20.01)	(20.37)
Final Confidence	54.57	64.72	62.90	62.63	65.76	59.74
	(34.42)	(31.26)	(29.41)	(30.91)	(28.26)	(27.92)
N participants	96	80	72	72	72	72
N matching groups	8	8	8	8	8	8
Top-half participants:						
IQ Quiz Score	10.13	10.10	10.26	10.32	10.66	10.03
	(1.78)	(1.35)	(1.69)	(1.36)	(1.47)	(1.59)
Initial Confidence	71.27	73.70	70.60	69.95	68.03	74.65
	(21.27)	(20.13)	(23.23)	(24.94)	(16.81)	(16.97)
Periods 1-12: Mean Confidence	73.99	82.49	77.08	75.75	72.76	74.37
	(21.63)	(15.43)	(17.24)	(21.27)	(19.16)	(18.65)
Final Confidence	78.29	86.20	79.54	78.24	78.83	76.43
	(26.11)	(17.64)	(19.92)	(19.55)	(24.56)	(24.23)
N participants	48	40	35	37	35	37
N matching groups	8	8	8	8	8	8
Bottom-half participants:						
IQ Quiz Score	6.42	6.38	5.78	6.43	6.27	6.11
	(1.54)	(1.73)	(2.08)	(1.58)	(2.13)	(1.62)
Initial Confidence	51.35	57.27	53.27	58.91	57.38	51.00
	(22.73)	(28.40)	(22.48)	(24.49)	(20.11)	(17.66)
Periods 1-12: Mean Confidence	44.94	64.19	65.43	68.79	58.30	51.00
	(20.18)	(22.62)	(23.08)	(21.33)	(18.39)	(14.48)
Final Confidence	30.85	43.25	47.16	46.11	53.41	42.09
	(23.74)	(26.88)	(28.41)	(32.31)	(26.11)	(19.56)
N participants	48	40	37	35	37	35
N matching groups	8	8	8	8	8	8

Notes: The table reports means. Standard deviations (calculated at the participant level) are in parentheses. Note that initial and final confidence reports were elicited absent any strategic motives in all treatments, whereas mean-reported confidence in periods 1–12 was elicited, including a strategic incentive to bluff in the *Strat* treatment and for the senders in *AsymNF* and *AsymF*. In *NStrat* and for the receivers in *AsymNF* and *AsymF*, there was no incentive to inflate reported confidence in periods 1–12.

Table A4: Robustness check: Regression results for round-by-round confidence updating (*NStrat* and Receivers in *AsymNF* and *AsymF* only).
Observations at floor or ceiling prior to updating are excluded.

	DV: Final (Non-Strategic) Confidence			
	(1)	(2)	(3)	(4)
Won	2.726***	2.726***	2.944***	2.752***
	(0.320)	(0.320)	(0.339)	(0.331)
Stayed out	1.626***	1.616***	1.146**	1.257**
	(0.330)	(0.545)	(0.525)	(0.593)
Lost	−2.939***	−2.939***	−2.874***	−2.782***
	(0.389)	(0.389)	(0.377)	(0.380)
Stayed Out X Feed/b Won		2.306	2.193	2.745
		(1.889)	(1.846)	(1.617)
Stayed Out X Feed/b Lost		−0.988	−0.922	−1.739**
		(0.926)	(0.878)	(0.767)
Opponent's Sign Confidence (cent.)			0.047***	0.113***
			(0.016)	(0.040)
Opponent's Sign Conf X Won				−0.108**
				(0.040)
Opponent's Sign Conf X Stayed Out				−0.078
				(0.046)
Opp's Sign Conf X Stayed Out X Feed/b Won				−0.042
				(0.056)
Opp's Sign Conf X Stayed Out X Feed/b Lost				0.093
				(0.060)
R^2	0.031	0.032	0.036	0.041
Observations	2305	2305	2305	2305

$^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$

Notes: OLS estimates. Absolute confidence update from one round to the next is the dependent variable (DV). Only observations from the *NStrat* treatment and from Receivers in the *AsymNF* and *AsymF* treatments are included. Furthermore, observations are excluded when the confidence from the previous round is already at the ceiling ($c = 100$) or floor ($c = 0$). The constant is suppressed. "Won," "Stayed Out," and "Lost" indicate whether a participant won, stayed out, or lost in the round preceding the confidence update. In Column (2), we include indicator variables that indicate whether a participant who stayed out in the *AsymF* treatment received feedback to have lost or to have won. In Column (3), we add the opponent's signal to a participant in the *AsymNF* and *AsymF* treatments received. In Column (4), we have the opponent's signal interact with a participant's experience prior to the participant's confidence update (i.e., with all variables of Column (2)). Standard errors clustered by 24 matching groups are reported in parentheses.

Table A5: Regression results for final confidence: Top vs. bottom-half on IQ Quiz.

	DV: Final (Non-Strategic) Confidence			
	(1)	(2)	(3)	(4)
Strat	8.299***	9.801**	5.825*	14.439**
	(2.336)	(4.354)	(3.019)	(5.494)
AsymNF / AsymF Senders	7.380**	13.757***	−0.966	2.120
	(2.784)	(4.637)	(3.709)	(5.809)
AsymNF Receivers	11.043***	19.912***	14.178***	27.996***
	(2.236)	(3.712)	(2.991)	(5.529)
AsymF Receivers	3.942	11.387**	8.605**	20.530***
	(3.067)	(4.826)	(3.740)	(6.030)
Strat x Top-half		−2.958		−9.750
		(6.549)		(6.970)
AsymNF / AsymF Senders x Top-half		−12.732**		−4.469
		(5.189)		(7.518)
AsymNF Receivers x Top-half		−17.954***		−22.594***
		(6.503)		(8.163)
AsymF Receivers x Top-half		−14.726**		−17.728**
		(6.290)		(7.045)
# Wins in Competition			4.173***	4.015***
			(0.529)	(0.817)
# Wins Unopposed			5.036***	4.978***
			(0.652)	(0.982)
# Stay Out			0.386	−0.759
			(0.585)	(0.790)
# Winning in Competition x Top-half				0.152
				(1.045)
# Winning Unopposed x Top-half				−0.294
				(1.177)
# Stay Out x Top-half				1.654
				(1.329)
Top-half in IQ Quiz	29.085***	38.710***	8.323***	14.565*
	(2.771)	(3.202)	(2.780)	(7.507)
Initial Confidence (centered)	0.444***	0.438***	0.322***	0.308***
	(0.056)	(0.057)	(0.062)	(0.064)
Constant	40.814***	35.991***	25.724***	22.832***
	(2.395)	(2.900)	(3.131)	(3.995)
R^2	0.468	0.480	0.583	0.592
Observations	464	464	464	464

$^{*}p < .10,\ ^{**}p < .05,\ ^{***}p < .01$

Notes: OLS estimates. Final confidence (scale from 0 to 100) is regressed on indicator variables for the different treatments. The *non-strategic* baseline is the omitted treatment. "#" variables are counts of how many times the indicated event occurred during the 12 rounds of the experiment. "Initial Confidence" is centered at the sample mean. Robust standard errors clustered by matching groups are reported in parentheses.

Table A6: Gender and round-by-round confidence updating (*NStrat* and Receivers).

	DV: Change in confidence (1)	(2)
Won	1.720***	1.530***
	(0.262)	(0.241)
Stayed out	0.322	0.326
	(0.616)	(0.724)
Lost	−3.899***	−3.741***
	(0.546)	(0.532)
Stayed Out X Feed/b Won	1.874	2.588*
	(1.689)	(1.271)
Stayed Out X Feed/b Lost	−0.730	−1.446
	(0.844)	(0.868)
Opponent's Sign Confidence (cent.)	0.033*	0.060
	(0.018)	(0.066)
Female	0.438	0.354
	(0.272)	(0.346)
Opponent's Sign Conf X Female	0.027	0.117
	(0.032)	(0.105)
Opponent's Sign Conf X Won		−0.052
		(0.069)
Opponent's Sign Conf X Won X Female		−0.132
		(0.109)
Opponent's Sign Conf X Stayed Out		−0.029
		(0.065)
Opponent's Sign Conf X Stayed Out X Female		−0.080
		(0.108)
Opp's Sign Conf X Stayed Out X Feed/b Won		0.037
		(0.052)
Opp's Sign Conf X Stayed Out X Feed/b Won X Female		−0.122
		(0.117)
Opp's Sign Conf X Stayed Out X Feed/b Lost		0.125
		(0.117)
Opp's Sign Conf X Stayed Out X Feed/b Lost X Female		−0.061
		(0.128)
R^2	0.028	0.035
Observations	2880	2880

$^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$

Notes: OLS estimates. Absolute confidence update from one round to the next is the dependent variable (DV). Only observations from the *NStrat* treatment and from Receivers in the *AsymNF* and *AsymF* treatments are included. "Won," "Stayed Out," and "Lost" indicate whether a participant won, stayed out, or lost in the round preceding the confidence update. "Stayed Out X Feed/b Won" and "Stayed Out X Feed/b Lost" capture whether a participant, who stayed out in the *AsymF* treatment, received feedback to have lost or to have won. Furthermore, we include an indicator for gender, the opponent's signaled confidence, and their interaction. In Column (2), we have the opponent's signaled confidence interact with a participants experience prior to the participant's confidence update. In addition, we include three-way interactions with gender. The constant is suppressed. Standard errors clustered by 24 matching groups are reported in parentheses.

Table A7: Regression results for final confidence: Gender.

	DV: Final (Non-Strategic) Confidence			
	(1)	(2)	(3)	(4)
Female	−4.231 (4.794)	−9.836 (7.610)	−2.312 (5.379)	−8.332 (8.529)
Tophalf		18.821*** (5.285)		17.618*** (4.750)
Female X Tophalf		10.883 (7.913)		11.368 (7.822)
Strategic	13.043*** (4.097)	9.537 (6.314)	8.120** (3.879)	2.969 (5.947)
Strategic X Female	−6.501 (5.903)	−3.300 (9.043)	1.161 (6.962)	6.115 (10.579)
Strategic X Top		5.059 (8.496)		8.252 (8.202)
Strategic X Fem X Top		−4.148 (9.352)		−7.575 (9.651)
AsymSender	10.608** (4.629)	8.964 (6.135)	8.213* (4.398)	5.024 (6.092)
AsymSender X Female	−4.880 (6.787)	4.305 (9.494)	−0.942 (6.913)	9.333 (9.800)
AsymSender X Top		2.082 (6.224)		5.084 (6.098)
AsymSender X Fem X Top		−19.100** (8.051)		−21.049** (8.125)
AsymNF Receiver	12.676** (5.167)	18.625*** (5.535)	11.246** (4.541)	15.232*** (5.542)
AsymNF Receiver X Female	−4.506 (7.689)	0.119 (9.755)	−2.250 (7.273)	2.234 (10.284)
AsymNF Receivers X Top-half		−8.715 (9.684)		−4.948 (10.225)
AsymNF Rec. X Fem X Top		−12.228 (10.528)		−11.411 (11.422)
AsymF Receiver	2.767 (5.870)	−1.086 (7.993)	1.909 (5.798)	−1.300 (7.759)
AsymF Receiver X Female	6.097 (7.531)	15.455 (10.671)	6.162 (7.929)	15.255 (10.719)
AsymF Receivers X Top-half		4.915 (10.248)		3.979 (9.254)
AsymF Rec. X Fem X Top		−17.370 (13.087)		−17.381 (11.300)
Initial Confidence (centered)			0.356*** (0.065)	0.350*** (0.061)
risk	0.280 (0.337)	0.193 (0.310)	0.072 (0.292)	−0.011 (0.263)
IQ Quiz Score	6.924*** (0.438)	4.066*** (0.559)	5.481*** (0.517)	2.556*** (0.575)
Constant	−2.198 (5.119)	13.137** (5.381)	10.645* (5.763)	27.199*** (6.297)
R^2	0.390	0.446	0.437	0.491
Observations	464	464	464	464

$^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$

Notes: OLS estimates. Final confidence (scale from 0 to 100) is regressed on indicator variables for the different treatments. The *NStrat* baseline treatment is omitted. "#" variables are counts of how many times the indicated event occurred during the 12 rounds of the experiment. "Initial Confidence" is centered at the sample mean. Robust standard errors clustered by matching groups are reported in parentheses. Testing for gender differences in the treatments, we find that final confidence for women is lower compared to men in the strategic treatment ($F(1, 31) = 9.98, p < 0.004$) and weakly so for Senders in the asymmetric treatment ($F(1, 31) = 3.45, p < 0.073$).

B Experimental Instructions

The original instructions were given in French, which is why the screenshots are in French. We provide the English translation below.

Explanations for this experiment

Welcome to this experiment!

From now on, it is **strictly forbidden to talk to the other participants**. If you have a question, ask the assistants. If you do not abide by this rule, we will have to exclude you from the experiment.

Today's experiment is comprised of six different parts. You will receive the instructions for each part at the beginning of each part, either on paper or directly on your computer screen.

You will find the instructions for the first part, a logic test, on the back of this sheet.

During this experiment you will make decisions that will affect your income as well as the income of other participants. Detailed rules concerning your income are explained in the instructions you will receive at the beginning of each part.

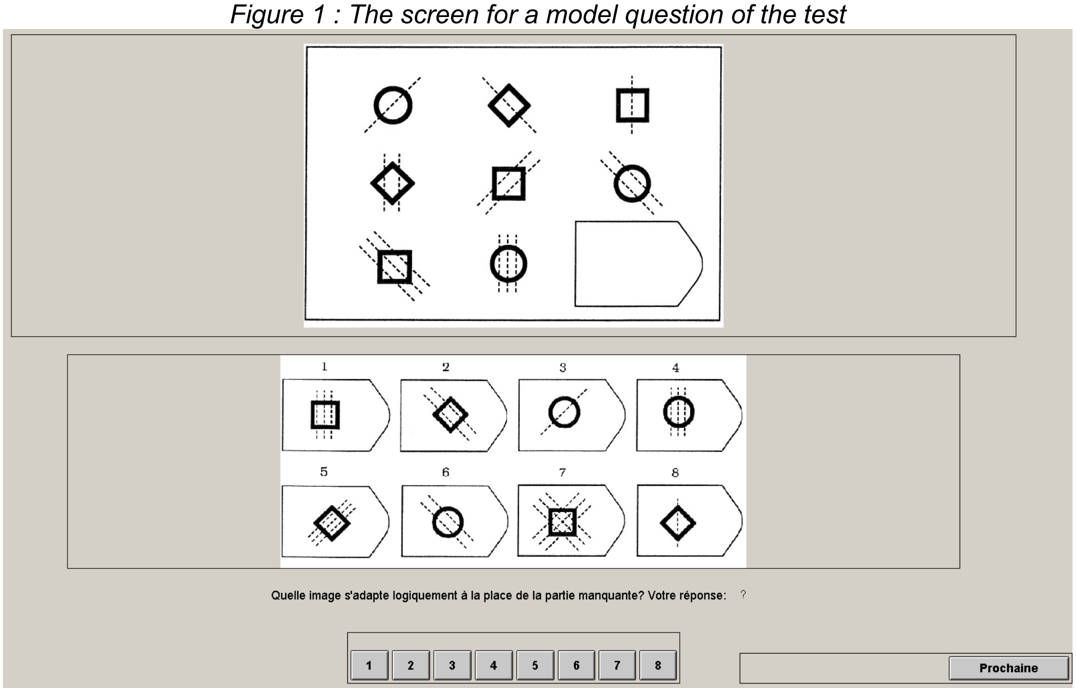
Unrelated to your decisions during the experiment, you will also receive a fixed amount of CHF 10.- for your participation.

1

Part 1: logic questions

You have 8 minutes to correctly answer as many of the questions that appear on the screen as possible. There are 15 questions in total. Each question refers to a sequence of pictures. The pictures are connected, but one element of the sequence is missing. Your task is to identify the missing element using the suggested answers. You can make your choice by clicking on the corresponding number at the bottom of the screen. You will do so with this screen:

Figure 1 : The screen for a model question of the test



You may skip a question and then return to it later on.

Once the time is up, the computer will calculate your score, which is the total amount of correct answers.

In part 3 of the experiment, you will participate in a game. To do so, you will be randomly partnered with another player and assigned a letter, either *A* or *B*. If you have a higher score on the logic test than the other player, you will be able to win more points.

2

Part 2: the evaluation

In this part, you can win points. Below you will find more details on how these points are assigned and the conversion rate into CHF.

I. Description of the evaluation

All of the participants have now finished with the questions, and we have determined each person's score.

You are randomly assigned to a group of 10 people, comprised of yourself and 9 other participants in this experiment. Your group will not change throughout the whole experiment.

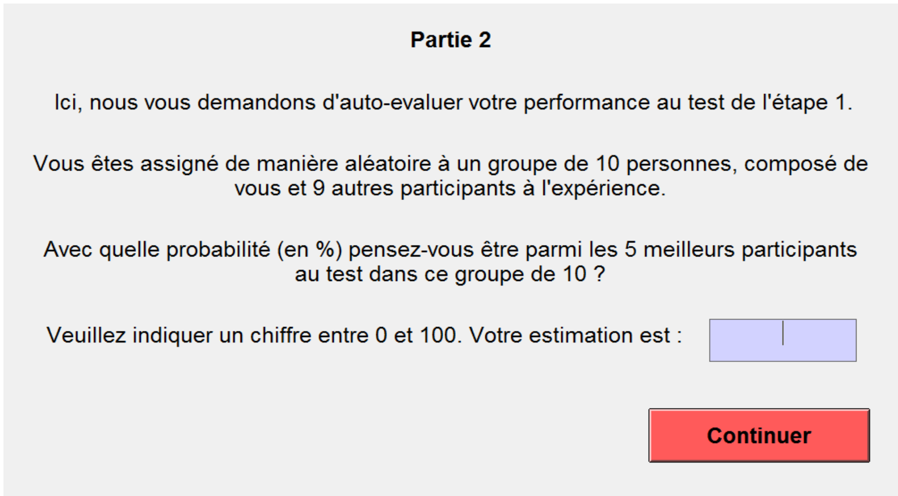
We have established a ranking of the participants in your group according to their score during the test in the previous stage. Now we are asking you to indicate how likely you believe it isthat you are among the best 50% (the best 5) participants in your group.

You can indicate this likelihood on a scale of 0 to 100%. Indicating 0% means that you are sure that you are not among the 5 best in your group, whereas indicating 100% means that you are sure that you are among the 5 best in your group. Likewise, 50% means that you believe that you are equally likely to be among the 5 best in your group or not to be among the 5 best in your group. The likelihood that you input will remain secret throughout the experiment.

You will win points for the accuracy of your estimation according to the rule described below.

If you wish, you may discover your real score and whether you are among the 5 best in your group at the end of the experiment. However, no one apart from yourself will ever see your score or your ranking.

You will enter your estimation with the following screen:



II. How do we calculate the points for the accuracy of your estimation?

You will win more points for your estimation the more accurate it is. The formula used is chosen in such a way that your expected winnings for the estimation are highest when you indicate your true belief. Indicating a value that differs from what you truly believe will reduce the number of points that you can receive for your estimation.

You start with 10 points. We subtract points according to the difference between your estimation and your true result, which is “1” if you are among the 5 best, and “0” if you are not.

The rule for calculating your points depends on the estimated likelihood that you entered on the screen. It corresponds to the following formula:
If you are among the 5 best, the points you win are:

$$\textit{points won} = 10 - 10 * [1-\frac{estimated\_likelihood}{100}]^2$$

If you are not among the 5 best, the points you win are:

$$\textit{points won} = 10 - 10 * [0-\frac{estimated\_likelihood}{100}]^2$$

The greater the difference between your estimation and the true result, the more points you will lose for your estimation. This loss is comparatively greater for large differences.

For example, if you declare that there is a 70% (0.7) likelihood that you are among the 5 best and according to your test result you are among the 5 best (your result is “1”), you are 0.3 away from the result, whereas if you are not among the 5 best (your result is “0”), you are 0.7 away from the result. As you can see in the formula above, the difference with the result is squared and multiplied by 10. Then we subtract this amount from the 10 points with which you started. So in the 70% example, if you are among the 5 best at the test, you receive $10-10(0.3)^2=9.1$. If you are not among the 5 best, you receive $10 – 10 (0.7)^2 = 5.1$.

In order to minimise the expected difference and to maximise the points that you will receive for your estimation, you should indicate your true belief.
For example, if you declare 70% but only believe that you have a 20% likelihood of being among the 5 best, you expect to win 9.1 points with a likelihood of 20% and 5.1 points with a likelihood of 80%.

The following examples illustrate the fact that the number of points that you can receive for your estimation is highest when you indicate your true belief. All of the numbers are used as examples and give no indication on which decisions to take.

Example 1:

You estimate the likelihood of being among the 5 best at 50% and you indicate 50%. If you indicate that there is a 50% chance that you are among the 5 best, there is always a difference of 0.5 with the result, and because that difference is squared, we will always subtract 10 times $(0.5)^2$, i.e. 2.5 points from the 10 points with which you started. This means that if you are among the 5 best, you receive $10 -10 (0.5)^2 = 7.5$ points for your estimation. If you are not among the 5 best, your winnings for the estimation are again $10 -10 (0.5)^2 = 7.5$.
If you believe that the likelihood of being among the 5 best is 50%, you should weight these two scores, 7.5 and 7.5, by the likelihood of realisation that you attribute to each one (50%). Your expected winnings for the estimation are 0.5 (7.5) + 0.5 (7.5) = 7.5.

You evaluate the likelihood of being among the best 5 at 50%, but you indicate 100%. If you indicate 100%, and you are indeed among the 5 best, you will receive 10 points for the estimation. But in the other case, if you are not among the 5 best, the difference is 1, so we subtract 10 times $(1)^2$ from the 10 points with which you started, and you will have 0 points. If you believe that the likelihood of being among the 5 best is 50%, you should weight these two scores, 10 and 0, by the likelihood of realisation that you attribute to each one (50%). The expected value of the subtracted amount is therefore 10 (0.5) + 0 (0.5) = 5. This gives

5

you an expected payoff of 5, which is smaller that the expected payoff if you declare your true belief of 50%, which is equal to 7.5 (as calculated above).

Example 2 :

You estimate the likelihood of being among the 5 best at 70% and you indicate 70%. If you indicate 70%, your gain for the estimation will be either 9.1 (if you are among the 5 best) or 5.1 (if you are not). You believe that there is a 70% chance that your winnings will be 9.1 and a 30% chance that your winnings will be 5.1. Therefore, the expected value of your winnings is 0.7 (9.1) + 0.3 (5.1) = 7.9.

You estimate the likelihood of being among the 5 best at 70% and you indicate 100%. Let us now assume that instead of indicating your belief of 70%, you indicate a different number. For example, you choose to declare 100%. This means that if you are among the 5 best, you will receive $10 - 10\,(0)^2 = 10$ points for your estimation. If you are not among the 5 best, you are 1 away from the true result and your winnings for the estimation will be $10 - 10\,(1)^2 = 0$. Since you expect to be among the 5 best with a likelihood of 70%, the expected value of your winnings for your estimation are 0.7 (10) = 7. This is smaller than 7.9, which is the expected value of your winnings if you indicate 70%.

You estimate the likelihood of being among the 5 best at 70% and you indicate 20%. The same is true if you indicate a number lower than your belief, for example 20%. If you are among the 5 best, your winnings for the estimation will be $10 - 10\,(0.8)^2 = 3.6$ points. If you are not among the 5 best, your winnings for the estimation will be $10 - 10\,(0.2)^2 = 9.6$. Since you actually expect to be among the 5 best with a likelihood of 70%, the expected value of your winnings for the estimation is 0.7 (3.6) + 0.3 (9.6) = 5.4. In other words it will again be lower than if you had indicated 70%.

The following table shows the expected results for several beliefs that you could have and for the estimations that you could give. As you can see, the expected winnings are highest when the estimation that you indicate is equal to your true belief (the cells on the diagonal that are indicated in bold).

Your indicated estimation (in %)	Your true belief (in %)										
	0	10	20	30	40	50	60	70	80	90	100
0	**10**	9	8	7	6	5	4	3	2	1	0
10	9.9	**9.1**	8.3	7.5	6.7	5.9	5.1	4.3	3.5	2.7	1.9
20	9.6	9	**8.4**	7.8	7.2	6.6	6	5.4	4.8	4.2	3.6
30	9.1	8.7	8.3	**7.9**	7.5	7.1	6.7	6.3	5.9	5.5	5.1
40	8.4	8.2	8	7.8	**7.6**	7.4	7.2	7	6.8	6.6	6.4
50	7.5	7.5	7.5	7.5	7.5	**7.5**	7.5	7.5	7.5	7.5	7.5
60	6.4	6.6	6.8	7	7.2	7.4	**7.6**	7.8	8	8.2	8.4
70	5.1	5.5	5.9	6.3	6.7	7.1	7.5	**7.9**	8.3	8.7	9.1
80	3.6	4.2	4.8	5.4	6	6.6	7.2	7.8	**8.4**	9	9.6
90	1.9	2.7	3.5	4.3	5.1	5.9	6.7	7.5	8.3	**9.1**	9.9
100	0	1	2	3	4	5	6	7	8	9	**10**

III. Conversion rate to CHF

The points won in this part of the experiment will be converted into CHF according to the following conversion rate:

1 point = 0.10 CHF

Note:
Parts in purple appear only in the baseline.
Parts in red appear only in the symmetric strategic treatment.
Parts in orange appear only in the baseline and the symmetric strategic treatment.
Parts in green appear only in the two asymmetric treatment.
Parts in blue appear only in the asymmetric treatment with feedback.

Part 3: the game

In this part you will take part in a game that allows you to win points. Below, you will find more details on how these points are attributed and the conversion of points into CHF.

I. Description of the game

In the game that you will play in this third part, there are two types of players: *A* and *B*. The roles of *A* or *B* have been randomly assigned.

You are a ***type A*** player.

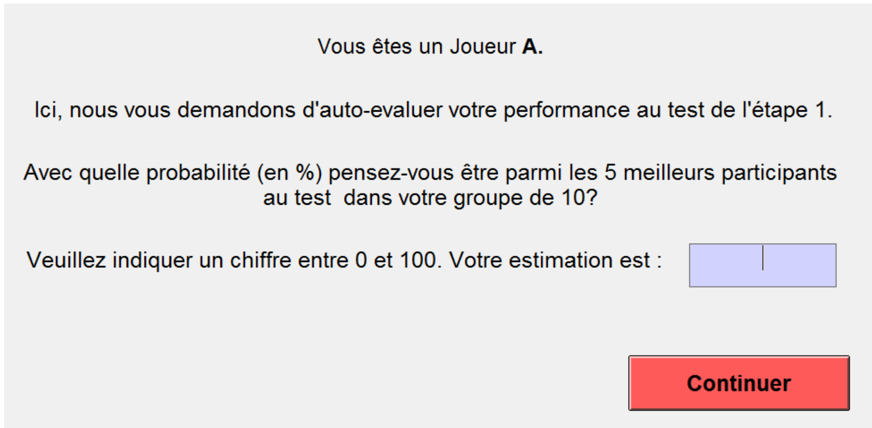
You will play the *A* role during the whole experiment, the same is true of the type *B* players.

You are assigned to the same group of 10 people to whom you had already been randomly assigned during the previous stage. To ease the explanation of the game, we will say that you are a type *A* player and that the other players in your group are type *B* players. Apart from the names, the roles of *A* and *B* are identical. The groups are such that each person in the group has a role different to the others. Therefore, as you are a type *A* player, **you are in a group of 10 people composed of yourself and 9 type *B* players.**

The game will be repeated for 12 turns with the same type *B* players who will always be a part of your group of 10. The game is played with two players and so each turn you will be randomly paired with a type *B* player of your group. Therefore, each turn, a type *B* player is assigned to you and will play with them. Each of the nine type *B* players in your group has the same likelihood of interacting with you, independently of which turn it is. It is therefore possible to meet the same player several times. However, the identities will stay hidden for all the participants during the whole of the experiment.

At the beginning of each turn, before you play the game with the type *B* player, we will ask you, as we did in the last stage, the likelihood with which you believe that you among the 5 best participants in your group in the logic test at the first stage of this experiment. We will ask the type *B* players the same question.

You will enter your estimation using the following screen:



We will pay you based on the precision of your estimation. The formula used for the allocation of points based on your precision is exactly the same as it was during the previous stage (part 2 of the experiment). Therefore, the more precise your estimation, the more points you will win. Furthermore, the formula used to calculate the points that you will win for the indication of your estimation is chosen in such a way that your expected winnings for the estimation are highest when you indicate what you truly believe. Indicating a different value to the one that you truly believe will reduce the number of points that you can receive for your estimation.

After the estimation, we will move on to the game in which you may also win points.

None of the players can see the estimation of being among the best 5 given by any other participant.

In the game, the test scores for you and for the *B* player with whom you are playing, taken at the previous stage, are compared.

The player who obtained the higher score will receive 10 points, the other will receive 0 points.

After each turn, you will be informed as to who had the higher score.

Before the game, you will see the *B* player's estimation of the likelihood of being among the 5 best in the group. The *B* player will also see your estimation of the likelihood of being among the 5 best players in the group.

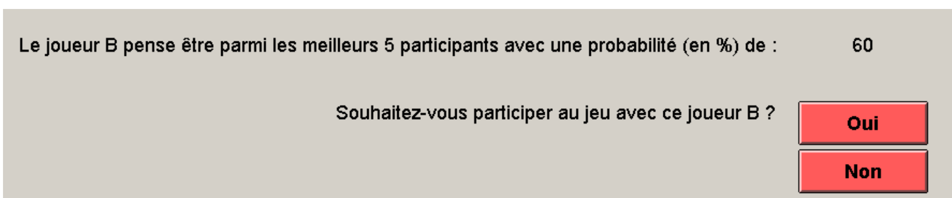
Each turn, after you have been shown the other player's estimation of being among the 5 best, you and the *B* player must decide if you want to play the game or not. If you and the *B* player both decide to play one against the other, then your test scores are compared.

If both you and the *B* player decide to play the game, your true test scores are compared. The player with the higher score wins 10 points. The player with the lower score wins 0 points. If both players have the same score, the winner is randomly selected.

If one of the two players decides not to play against the other, that player wins 5.5 points. If both players decide not to play, they both win 5.5 points.

If one player decides to play against the other, but the second player decides not to play, the player who decided to participate in the game automatically wins 10 points.

After having seen the other's estimation of being within the 5 best, you and the *B* player make the decision of playing the game or not by using this screen:



If both players decide to play the game then at the end of each turn you and the *B* player both find out who had the higher score.

Before the game, as you are an *A* player, you will not see the estimation given by the *B* player. However, the *B* player will see your estimation of the likelihood of being among the 5 best participants in your group.

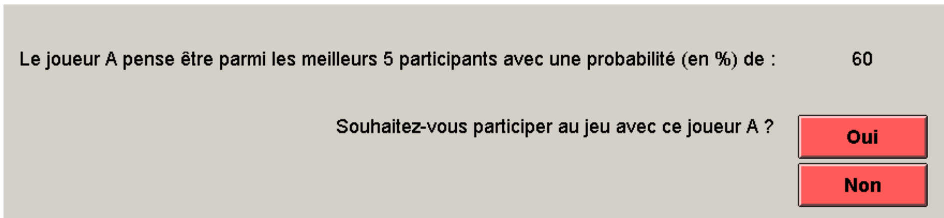
Each turn, after having seen your estimation of being among the 5 best, the *B* player must decide whether or not they wish to participate in the game with you. If the *B* player decides to play against you, then during the game the test scores of yourself and of the *B* player are compared.

If the *B* player decides to play against you, the true test scores are compared. The player with the higher score receives 10 points. The player with the lower score receives 0 points. If both players have the same score, the winner is chosen randomly.

If the *B* player decides not to play against you, you win 10 points and the *B* player wins 5.5 points.

10

After having been informed of your estimation of the likelihood of being among the 5 best, the *B* player decides whether or not they wish to participate in the game using the following screen:



At the end of each turn, you will learn whether or not the *B* player decided to play against you. If they did, both the *B* player and yourself are informed as to who had the higher score.

If the event that the *B* player decided not to play against you, the *B* player will learn who had the highest score regardless. You will not be given this information in this situation.

II. Remuneration and rate of conversion of points into CHF

You will collect points for your results in the game and for the precision of the estimation of your test result in each of the twelve turns of this stage. **The computer will randomly choose one of the twelve turns.** You will be paid according to your result during the turn chosen by the computer. That means that you will receive points for the precision of the estimation of your test result as well as the points won during the game in the same turn. Each of the twelve turns has the same likelihood of being chosen by the computer.

The points won during the turn randomly chosen by the computer will be converted into CHF using the following conversion rate:

1 point = 1.20 CHF

11

III. Have you understood properly?

Before starting, we want to be sure that you and the other participants have understood the decisions that you may take. For this, please look at the following questions and answers.

Question 1.
Assume that you have a higher test score than the *B* player. How many points will you receive and how many will the *B* player receive?

Answer: You win 10 points and the *B* player wins 0 points.

Question 2.
Assume that you have a lower test score than the *B* player. How many points will you receive and how many will the *B* player receive?

Answer: You win 0 points and the *B* player wins 10 points.

Question 1.
a) Which information will you receive at the beginning of each game turn?

Answer: You will be informed of the*B* player's estimation of being among the 5 best participants in their group of 10.

b) Which information will the *B* player receive at the beginning of each game turn?

Answer: The *B* player will be informed of the estimation given by yourself, the *A* player, of being among the 5 best participants in your group of 10.

c) Which decision will you takebefore the game?

Answer: Each turn, you will decide whether you wish to play with the *B* player or not.

d) Which decision will the *B* player make before the game?

Answer: Each turn, the *B* player decides whether or not they wish to play the game with you.

Question 2.
Assume that you have a higher test score than the *B* player.

a) How many points will you receive and how many points will the *B* player receive if both yourself and the *B* player decide to play against one another?

Answer: You win 10 points and the *B* player wins 0 points.

b) How many points will you receive and how many points will the *B* player receive if you decide to play the game, but the *B* player decides not to play against you?

Answer: You win 10 points and the *B* player wins 5.5 points.

c) How many points will you receive and how many points will the *B* player receive if both yourself and the *B* player decide not to play the game?

Answer: you win 5.5 points and the *B* player wins 5.5 points.

d) How many points will you receive and how many points will the *B* player receive if you decide not to play against the *B* player, but the *B* player decides to play the game?
Answer: You win 5.5 points and the *B* player wins 10 points.

Question 3.
Assume that you have a lower test score than the *B* player.

a) How many points will you receive and how many points will the *B* player receive if both yourself and the *B* player decide to play against one another?

Answer: You win 0 points and the *B* player wins 10 points.

b) How many points will you receive and how many points will the B player receive if you decide to play the game, but the *B* player decides not to play against you?

Answer: You win 10 points and the *B* player wins 5.5 points.

c) How many points will you receive and how many points will the *B* player receive if both yourself and the *B* player decide not to play the game?

Answer: you win 5.5 points and the *B* player wins 5.5 points.

d) How many points will you receive and how many points will the *B* player receive if you decide not to play against the *B* player, but the *B* player decides to play the game?
Answer: You win 5.5 points and the B player wins 10 points.

Question 1.
a) Which information will the *B* player receive at the beginning of each game turn?

Answer: The *B* player will be informed of the *A* player's estimation of being among the 5 best participants in their group of 10.

b) Which decision will the *B* player take before the game?

Answer: Each turn, the *B* player will decide whether or not they wish to play the game with the *A* player.

Question 2.
Assume that you have a higher test score than the *B* player.

a) How many points will you receive and how many points will the *B* player receive if the *B* player decides to play against you?

Answer: You win 10 points and the *B* player wins 0 points.

b) How many points will you receive and how many points will the *B* player receive if the *B* player decides not to play against you?

Answer: You win 10 points and the *B* player wins 5.5 points.

Question 3.
Assume that you have a lower test score than the *B* player.

a) How many points will you receive and how many points will the *B* player receive if the *B* player decides to play against you?

13

Answer: You win 0 points and the *B* player wins 10 points.

b) How many points will you receive and how many points will the B player receive if the *B* player decides not to play against you?

Answer: You win 10 points and the *B* player wins 5.5 points.

C Theoretical Model

We present a simple Sender-Receiver game of entry deterrence with lying costs. The model uses payoffs similar to those used in the experiment, despite clear differences in the information structure. The model clearly shows that players have an incentive to over-report their type (top half or bottom half) in equilibrium.

Each member of a group of Senders of mass 1 and a group of Receivers of mass 1 undertake an identical task and obtain an outcome or score as a result of their endeavor. The median score is the same in both groups. Since group formation is exogenous, from the player's perspective, being above or below the median can be seen as a random draw. To simplify the exposition, let us assume that, with probability 0.5, a player is a high type θ_1 with a score above the median and, with probability 0.5, she is a low type θ_0 with a score below the median.

Players do not know their type with certainty but only observe a signal of their type.[21] Based on this signal, (1) Senders have to decide what signal they want to report to Receivers, and (2) Receivers have to decide whether they want to compete against a Sender, given their own signal and the Sender's message.

In the following, we show that rational Senders have an incentive to over-report their signal even if honesty is rewarded. Rational Receivers have an incentive to opt out of the competition whenever they receive a high signal.

C.1 Information

The signal s is either high or low, say 1 or 0. The signal is informative, so the θ_0-type player is more likely to observe $s = 0$, and the θ_1-type player is more likely to observe $s = 1$. Let p be the precision of the signal: $p = Pr(s = i \mid \theta_i)$. A player θ_i observes a signal $s = j$ for $i, j = 0, 1$ and $i \neq j$ by mistake with probability $1 - p$. Informativeness of the signal implies that p is greater than $1/2$.

After Nature has assigned types and signals realize, player infers her own type θ_i by applying Bayes' rule:

$$Pr(\theta_i \mid s = i) = \frac{Pr(s = i \mid \theta_i)Pr(\theta_i)}{Pr(s = i \mid \theta_i)Pr(\theta_i) + Pr(s = i \mid \theta_j)Pr(\theta_j)} = p,$$
$$Pr(\theta_i \mid s = j) = \frac{Pr(s = j \mid \theta_i)Pr(\theta_i)}{Pr(s = j \mid \theta_i)Pr(\theta_i) + Pr(s = j \mid \theta_j)Pr(\theta_j)} = 1 - p.$$

[21] In our experimental setup, this signal corresponds to the information generated in the previous round of play.

For the analysis below, note that the unconditional probability of the signal being equal to 0 or 1 corresponds to

$$Pr(s=0) = Pr(s=0 \mid \theta_0)Pr(\theta_0) + Pr(s=0 \mid \theta_1)Pr(\theta_1) = \frac{1}{2},$$
$$Pr(s=1) = Pr(s=1 \mid \theta_1)Pr(\theta_1) + Pr(s=1 \mid \theta_0)Pr(\theta_0) = \frac{1}{2}.$$

To save on notation in the following, we denote a high signal $s = 1$ as s_1 and a low signal $s = 0$ as s_0. Similarly, a high type is denoted θ_1, and a low type is denoted θ_0.

C.2 The Sender-Receiver Game

Two players, a Sender and a Receiver, play the following game:

1. Nature assigns types with probability $\frac{1}{2}$ and players observe their realized signals s.

2. The Sender chooses to report a message to a randomly chosen Receiver. The message is denoted by m, and $m \in \{0, 1\}$.

3. The Receiver observes the message m, makes an inference about the signal of the sender and, knowing her own signal s^R, decides whether to enter into a competition with the Sender or not (play IN or OUT).

4. If a competition takes place, the player with the higher θ type wins b. In the case of a tie, a winner is drawn at random such that the expected payoff is $b/2$. The loser in a competition gets a payoff of zero. The Receiver always has the option to avoid the comparison with the Sender and collect the payoff O^R. To replicate the experimental design, we set $O^R = \frac{b}{2} + \epsilon$.

5. The Sender's payoff consists of the competition payoff described above plus an honesty payoff reflecting the degree of accuracy of her message. The calculation of the honesty payoff is described below.

C.3 Cost of lying

Senders are rewarded for correctly indicating their θ type. The closest approximation to their type is indicated by their signal, although with some imprecision. Like in the experimental design, we introduce a payoff (denoted by H) for being honest and correct. Let $m = i$, with $i = 0, 1$, be the reported type and let b be a positive value. Then Senders can maximize their honesty payoff by reporting $m = s$ (i.e., a message equal to the observed signal). If one observes $s = 0$ and reports $m = 0$, the Honesty payoff is $H = p[b - b(0-0)^2] + (1-p)[b - b(1-0)^2] = pb$. The payoff would be identical if someone had observed $s = 1$ and reported $m = 1$: $H = p[b - b(1-1)^2] + (1-p)[b - b(0-1)^2] = pb$. If one observes $s = 0$ but reports $m = 1$, then the honesty payoff is $H = p[b - b(0-1)^2] + (1-p)[b - b(1-1)^2] = (1-p)b$. Table A8 summarizes the Honesty payoffs according to the signal and the report.

Table A8: Honesty payoffs.

	$\mathbf{m = 0}$	$\mathbf{m = 1}$
$\mathbf{s = 0}$	pb	$(1-p)b$
$\mathbf{s = 1}$	$(1-p)b$	pb

C.4 Equilibrium Analysis: Honest Equilibrium

We look for a perfect Bayesian equilibrium for the Sender-Receiver game, where the Senders always report their observed signal. Upon observing the message m and correctly inferring that the Sender has observed a signal $s^S = m$, the Receiver's decision depends on the Receiver's signal s^R. Table A9 summarizes the expected payoffs when the Receiver with signal s^R chooses to play IN [i.e., $\pi^R(IN \mid s^R, s^S)$].

To illustrate how the payoffs are calculated, consider the first quadrant of Table A9. The Receiver has signal $s^R = 0$ and receives a message $m = 0$ from a random Sender. Because we assume that honest reports are an equilibrium strategy for the Sender, we expect the Receiver to believe that the Sender has signal $s^S = 0$. For this Sender, the expected payoff from playing IN is

$$\pi^R(IN \mid s_0^R, s_0^S) = pp\frac{b}{2} + p(1-p)0 + (1-p)pb + (1-p)(1-p)\frac{b}{2} = \frac{b}{2}.$$

The first term $pp\frac{b}{2}$ is the probability that the Receiver is of type θ_0 times the probability that the Sender is of type θ_0 times the payoff when there is a tie ($b/2$). The other terms and payoffs are calculated analogously.

Table A9: Receiver payoff from playing IN, $\pi^R(IN \mid s^R, s^S)$.

	Receiver has signal 0 $s^R = 0$	**Receiver has signal 1** $s^R = 1$
Sender $s^S = m = 0$	$\frac{b}{2}$	pb
Sender $s^S = m = 1$	$b(1-p)$	$\frac{b}{2}$

The optimal action for a Receiver given her beliefs about the Sender type depends on the comparison between the expected payoff from competing and the outside option O^R. Since $O^R = \frac{b}{2} + \epsilon$, the best response of a Receiver can be summarized as follows:

LEMMA 1: If Senders report their signals truthfully, then a Receiver with signal $s^R = 0$ will always choose OUT; a Receiver with signal $s^R = 1$ will stay OUT after observing $m = s = 1$ and will play IN after observing $m = s = 0$ provided $\epsilon < b(p - \frac{1}{2})$.

We are left to consider whether the message-strategy of the Sender given the subsequent behavior of the Receivers maximizes payoff. A Sender with signal $s^S = 1$ who reports $m = 1$ expects to make $\pi^S(m = S \mid s^S = 1) = pb + b$, where the first term corresponds to the Honesty payoff and the second term to the expected payoff from competition. She has no incentive to report a message lower than the signal since this will lower both her honesty payoff and her expected payoff from competition. A Sender with signal $s^S = 0$ who reports $m = 0$ expects to make

$$\begin{aligned}
\pi^S(m = S \mid s_0^S) &= H(m = s) + Pr\left(s_0^R\right) b + Pr\left(s_1^R\right) \\
&\quad \times \left\{ Pr\left(s_0^S \mid \theta_0^S\right) \left[Pr\left(s_1^R \mid \theta_0^R\right) \times b/2 + Pr\left(s_1^R \mid \theta_1^R\right) \times 0 \right] \right. \\
&\quad \left. + Pr\left(s_0^S \mid \theta_0^S\right) \left[Pr\left(s_1^R \mid \theta_0^R\right) \times b + Pr\left(s_1^R \mid \theta_1^R\right) \times b/2 \right] \right\} \\
&= b + p\frac{b}{2}.
\end{aligned}$$

If a Sender with signal $s = 0$ deviates and reports $m = 1$, her payoff from honesty decreases to $(1 - p)b$, but she will be able to deter all opponents and gain b. That is, her expected payoff is $\pi^S(m = S \mid s_0^S) = (1 - p)b + b$. The deviation for a Sender with $s^S = 0$ is thus profitable if

$$(1 - p)b + b > b + p\frac{b}{2},$$

which holds if $p < \frac{2}{3}$. That is, for $p < \frac{2}{3}$, a Sender with a low signal has an incentive to over-report and send message $m = 1$. By contrast, if p is higher (i.e., if the signal is very precise), then over-reporting is not profitable because it entails a large loss in terms of the honesty payoff. That is, we get a separating equilibrium with honest reporting of signals if the signal is sufficiently informative. This leads to the following lemma:

LEMMA 2: If $p > \frac{2}{3}$, the equilibrium strategy for a Sender is to choose $m = s^S$ (i.e., to report her signal truthfully). Receivers with $s^R = 0$ will never compete, irrespective of the message received, whereas Receivers with signal $s^R = 1$ will choose to play IN after receiving the message $m = 0$ and OUT otherwise.

C.5 Equilibrium Analysis: Over-Reporting Equilibrium

We now assume that $\frac{1}{2} < p < \frac{2}{3}$. That is, the signal is informative but not sufficiently precise. We look for a hybrid equilibrium where the Sender with signal $s^S = 0$ randomizes between reporting $m = 0$ and $m = 1$. Let r $((1-r))$ be the beliefs held by a Receiver after seeing the message $m = 1$ that the Sender has signal $s^S = 1$ $(s^S = 0)$. Let x be the probability that a Receiver with signal $s^R = 1$ plays IN after seeing $m = 1$. We proceed in steps:

Step 1: If a Receiver observes the message $m = 0$ she correctly infers that the Sender has signal $s^S = 0$. According to the previous payoff analysis, we conclude that the Receiver with signal $s^R = 0$ will play OUT whereas the Receiver with signal $s^R = 1$ will play IN.

A Sender with signal $s^S = 1$ will report her signal truthfully, as discussed above.

Step 2: A Sender with signal $s^S = 0$ may want to randomize between sending $m = 0$ and $m = 1$. The payoff from sending $m = 0$ is pb for being honest, plus the expected payoff $\frac{1}{2}b$ from being matched with a Receiver with signal $s^R = 0$ who chooses to stay OUT, plus the expected payoff from being matched with a Receiver of type $s^R = 1$ which is $\frac{1}{2}[pp0 + p(1-p)\frac{b}{2} + (1-p)p\frac{b}{2} + (1-p)(1-p)b] = \frac{1}{2}[b(1-p)]$. Overall, the payoff from sending $m = 0$ when $s^S = 0$ is

$$\pi^S\left(m = 0 \mid s_0^S\right) = b + \frac{1}{2}bp.$$

The expected payoff for sending $m = 1$ is

$$\pi^S\left(m = 1 \mid s_0^S\right) = (1-p)b + \frac{1}{2}b + \frac{1}{2}\Big\{x[b(1-p)] + (1-x)b\Big\},$$

where x is the probability that a Receiver with a high signal $s^R = 1$ plays IN after receiving a high message. The Sender with signal $s^S = 0$ is indifferent between sending $m = 0$ and $m = 1$ if the Receiver with signal $s^R = 1$ plays IN with probability $x = \frac{2}{p} - 3$. Since $x \in [0, 1]$, we must have $\frac{1}{2} \leq p \leq \frac{2}{3}$. This finding is consistent with the previous result: when $p > 2/3$, the honest equilibrium arises.

Step 3: The condition for a Receiver with signal $s^R = 1$ to be indifferent between playing IN or OUT after seeing $m = 1$ is

$$r\frac{b}{2} + (1-r)bp = \frac{b}{2} + \epsilon;$$

that is, if

$$r = \frac{b[p - \frac{1}{2}] - \epsilon}{b[p - \frac{1}{2}]}.$$

As explained above, r is the belief held by a Receiver that the sender has signal $s^S = 1$ after seeing the message $m = 1$.

Step 4: Finally, beliefs must be derived by using Bayes' rule and the strategy of the Sender in equilibrium. Let λ be the probability that a Sender with signal $s^S = 0$ sends a high message. According to Bayes' rule, when a Receiver observes the message $m = 1$, she infers that the Sender has signal $s^S = 1$ with probability

$$Pr(s_1^S \mid m = 1) = \tfrac{1}{1+\lambda}.$$

By equating r to $\frac{1}{1+\lambda}$, we find that the Sender with signal $s^S = 0$ reports the message $m = 1$ with probability $\lambda = \frac{\epsilon}{b(p-\frac{1}{2})-\epsilon}$.

We can now sum up the perfect Bayesian equilibrium when Senders with $s^S = 0$ over-report with positive probability:

LEMMA 3: For $1/2 < p < 2/3$, there exists an equilibrium such that

i. Senders with $s^S = 1$ always report $m = 1$.

ii. Senders with $s^S = 0$ report $m = 1$ with probability $\lambda = \frac{\epsilon}{b(p-\frac{1}{2})-\epsilon}$.

iii. Receivers with $s^R = 0$ always choose to stay OUT.

iv. Receivers with $s^R = 1$ always choose to play IN after $m = 0$ and play IN after seeing $m = 1$ with probability $x = \frac{2}{p} - 3$.

This simple model illustrates that Senders have an incentive to overstate their type if the players' signals about their type are not sufficiently precise. Such behavior decreases the probability that Receivers enter the competition compared with the equilibrium with honest reporting. This can be interpreted as an incentive to bluff, which would increase the deterrence of opponents. This conclusion is consistent with the behavior observed in the experiment.