TI 2011-151/1
Tinbergen Institute Discussion Paper

# Self-confidence and Strategic Deterrence

*Gary Charness*[1]
*Aldo Rustichini*[2]
*Jeroen van de Ven*[3]

[1] *University of California at Santa Barbara;*
[2] *University of Minnesota;*
[3] *University of Amsterdam, and Tinbergen Institute.*

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031



Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

# Self-confidence and strategic deterrence

Gary Charness, Aldo Rustichini, and Jeroen van de Ven

July 26, 2011

**Abstract:**

We examine factors that may contribute to 'overconfidence' in relative ability on an intelligence test. We test experimentally for evidence of self-esteem concerns and instrumental strategic concerns. Errors in Bayesian updating are rare when the information does not involve own relative ability, but far more common when it does, suggesting self-esteem issues. There is also strong evidence that males state higher levels of confidence in relative ability when this precedes a tournament; as entry is predicted by relative confidence, this can be an effective deterrent. Inflating confidence can be part of an equilibrium strategy, providing a rationale for strategic overconfidence.

**Keywords:** Self-confidence, overconfidence, strategic deterrence, unconscious behavior, self-deception, experiment

**JEL Classifications:** A12, C91, D03, D82

**Contact:** Gary Charness, University of California at Santa Barbara, charness@econ.ucsb.edu, Aldo Rustichini, University of Minnesota, arust@econ.umn.edu, Jeroen van de Ven, University of Amsterdam, j.vandeven@uva.nl.

# 1. Introduction

Beliefs about one's abilities are an important ingredient to many decisions, including making career choices, undertaking enterprises, and taking risks. Many studies in psychology and economics support the claim that people are overconfident in their ability (e.g., Svenson, 1981; Dunning, Meyerowitz, and Holzberg, 1989).[1] Often this evidence comes from verbal statements by people on their confidence in their relative ability, but some studies also show evidence of overconfidence in choice behavior (e.g., Hoelzl and Rustichini, 2005).[2] Such overconfidence can have negative consequences for people's choices and their corresponding economic outcomes.[3] So an over-arching question concerns the roots of such overconfidence in relative ability and the corresponding benefits that might explain the persistence of the phenomenon.

There are at least two main candidates for personal benefits from such overconfidence. First, one may receive consumption value (ego utility, in Koszegi, 2006) from the belief that one is talented. In this view, people feel better with a favorable self-perception, even at the cost of being overconfident when reality intrudes. Second, one may value the belief of others about one's ability, either for non-instrumental or instrumental purposes. Presumably one likes good (but not bad) information to be known. In the non-instrumental case, projecting high levels of confidence may enable one to simply feel the warm glow of social approval. In the alternative instrumental case, the social signal of confidence may also have the value of favorably modify-

[1] Moore and Healy (2008) mention a taxonomy of overconfidence, consisting of "(1) overestimation of one's actual performance, (2) overplacement of one's performance relative to others, and (3) excessive precision in one's beliefs." In this paper, we primarily consider the second of these categories, and hereafter refer to this as "overconfidence".

[2] It is also found that overconfidence can turn into underconfidence when task difficulty increases (Hoelzl and Rustichini, 2005; Moore and Cain, 2007).

[3] For example, Dohmen and Falk (2006) find that overconfident individuals are more likely to select themselves into a tournament contract, and Malmendier and Tate (2008) find that confident CEOs are more prone to take value-destroying merger decisions. In other field studies it is found that investors trade too much (Odean, 1999) and consumers overestimate their future attendance of health clubs (DellaVigna and Malmendier, 2006); see DellaVigna (2009) for more examples.

ing the behavior of others. For example, a lawyer or salesperson might well need to project confidence to prospective clients or customers in order to be hired or to make a sale. Thus overconfidence may have a strategic role, increasing one's own utility in a game or competition. A novel contribution of our paper is to investigate whether apparent overconfidence can in fact be strategic, either on a conscious or unconscious level.

Our experiments investigate overconfidence relative to others. First, we use an incentive-compatible mechanism to elicit confidence in one's relative ability in a cognition task, in order to detect signs of the presence of overconfidence. Next, we consider an updating task where one receives a negative (but noisy) signal. We then compare behavior in an environment in which people receive feedback about their performance on a mental-ability task and in an isomorphic setting where this feedback relates to an abstract question. This comparison provides evidence to test whether people are generally unable to process information effectively, or whether there is a self-serving bias when one's own performance or ability is at stake.

In additional treatments, we test to see if reported confidence is sensitive to social saliency. To study this, we manipulate saliency of the social signal (reported confidence shown to another participant) and vary whether there could be a strategic motivation to report high confidence. Each person knows that one's stated confidence will be reported to another person with whom one is paired (this is common information). When the confidence is simply reported to another person with no payoff consequences, this provides a test of whether people value a social signal in an anonymous environment. Instead, in the strategic setting there are payoff consequences: one person in a pair is automatically entered into a tournament, while the other decides whether to enter the tournament or to choose an outside payment option. Since winning the tournament is much better than losing, the first person has strategic reasons to deter the other person from entering. Thus, stating a higher level of confidence may be an effective deterrent (and in fact it is in our data). Although misreporting one's beliefs is costly, strategic deterrence can be part of an equilibrium in which people (consciously or unconsciously) inflate their reported confidence.

To summarize our findings, first we find that errors in updating about own ability are far more likely than the rare errors in updating about an abstract scenario. Since the different conditions are equivalent from the point of view of a Bayesian observer, if the conditions lead to different beliefs something other than Bayesian updating must be taking place. This indicates that people are fairly competent at processing neutral information in the task we chose, but not personal information; this suggests there may be consumption value in believing that one is better than average. Second, the voluntary-entry treatment provides evidence that stated confidence (for males but not females) is higher in a strategic (competitive) environment relative to the social non-strategic environment or the baseline.[4] There is considerable value in reporting a higher confidence level, as the likelihood of entry by a person with discretion about entering the tournament is much higher when that person's stated confidence is as high as the other person's than when it is lower; so the higher stated confidence levels act as a deterrent.

Males inflate their stated confidence levels in this competitive environment to the same extent, regardless of whether the other person can be deterred from entering the tournament. We know that debriefed participants did not report a conscious awareness of inflating their stated confidence levels, nor was there any evidence of any sort of calculations. Thus the fact that in-

[4] We find that the mean reported estimate in the baseline treatment about the likelihood that one is above average is 63.4 percent. Whether this is *per se* evidence of overconfidence is controversial (Benoit and Dubra, 2011; Burks, Carpenter, Goette and Rustichini, 2010). In any case, we do not find any effect of a pure social signal, as the mean reported estimate in the social-information treatment about the likelihood that one is above average is almost identical, at 63.5 percent.

flating is common to players in both roles could be evidence of unconscious decision-making. Competition *per se* might lead to an automatic response in behavior. It might well be that people even need to be unaware of their bias in order to influence themselves or others successfully. To the extent that overconfidence functions at an unconscious level, self-deception seems an important component of the process.

Finally, we see that females are significantly less likely to enter the tournament. However, this does not appear to be driven by shying away from competition (Niederle and Vesterlund, 2007), but instead by confidence level, as there is no significant difference in entry rates when we control for confidence. The difference in confidence across gender is only significant in the strategic environment. In fact, the difference in confidence across gender only manifests for those people who choose to enter the tournament, with little difference in stated confidence levels for men and women who choose not to enter into the competition.

The remainder of this paper is structured as follows. In section 2, we provide a review of the literature, and we describe our hypotheses and our experimental design in section 3. We present our experimental results in section 4, and we discuss these in section 5. We conclude in section 6.

## 2. Background and literature review

Social psychology has long considered the issues of self-esteem, overconfidence, and self-deception. Baumeister (1998) provides an extensive review of the overconfidence phenomenon. Further evidence and discussion on the topic of self-esteem can be found in Leary, Tambor, Terdal, and Downs (1995) and Leary (1999), where image concerns lead to a selective demand for information. Berglas and Jones (1978) and Kolditz and Arkin (1982) also study how self-handicapping is related to social saliency. Kolditz and Arkin (1982) find that subjects take performance-impoverishing drugs after receiving positive feedback about their past performance when their choice of drugs is visible to the experimenter. However, when subjects choose whether or not to take the performance-impoverishing drugs in private, no subjects take them. This suggests that performance/confidence is a social signal.

Rabin and Schrag (1999) provide a model of confirmatory bias, where people misinterpret new information as supporting previously held views; in this model such confirmatory bias induces overconfidence. An agent may come to believe with near certainty in a false hypothesis, even though he or she receives an unlimited amount of information. Koszegi (2006) provides a formal economic model of overconfidence and ego utility, in which an agent derives internal benefits from positive views about his or her ability. The mechanism in this model is that each person receives an initial signal about own ability and can seek information if desired.

A number of very recent or contemporaneous papers examine overconfidence; to date only one of these has been published. Eil and Rao (2011) study how people process and acquire objective information regarding intelligence and beauty. They rank 10 people in a group, and elicit the complete belief distribution over all possible ranks; participants receive incomplete but truthful feedback. Each person's rank was compared bi-laterally to an anonymous and random-chosen participant, with each then told whether they rank high or lower. People are found to respond much more to positive feedback than to negative feedback. Negative feedback leads to less predictable updating behavior and also a dislike for acquiring new information. When the information is about a non-personal phenomenon, updating and information acquisition were unbiased.

The advantage of their set-up is that one gets more precise information; on the other hand, this also makes it much more complicated for participants. Even in the control treatment with a neutral task, where self-image plays no role, updating seems to be quite noisy and errors appear to be common.[5] The fact that updating is unbiased in the neutral task may be due to the fact all subjects start from the same symmetric (uniform) distribution in that case: each rank is equally likely at the start. The advantage of our simple set up is that participants make very few errors on the neutral machine task, and even people with very low confidence make very few errors. We also ensured that the distributions of the priors are similar in both tasks, so that differences in updating cannot be due to different prior distributions. Moreover, strategic issues are neither part of their study, nor part of any of the other studies below.

Mobius, Niederle, Niehaus, and Rosenblat (2005-2011) study how subjects respond to noisy feedback about their performance in an IQ test; people receive signals about their relative rank that are correct with 75% probability. They find that subjects are conservative (not responding enough to new information) and asymmetric (reacting more to positive feedback than to negative feedback); women are more conservative than men, but not more asymmetric. High-confidence women have a higher value for information than men. They develop a model of biased information processing inspired by the anticipatory utility model of Brunnermeier and Parker (2005), where these seemingly related phenomena naturally arise. They do not compare updating to a neutral case where self-image concerns are absent.

Ertac (2008) gives feedback to subjects in two different environments.[6] In one case, feedback is about urns (non-performance related). In the other case, feedback is about performance in an algebra and verbal test. She finds no systematic bias in updating in the case of a neu-

[5] There is also a possible alternative interpretation of their results, in that people with high confidence on the intelligence tests are smart, and better at updating. Since they are more likely to get positive feedback, the results could reflect the fact that smarter people are better at updating. However, since they find similar results in the case of beauty, this alternative interpretation is only correct if better-looking people are better at updating.

[6] Our study was conducted independently, as we were unaware of this unpublished work when we designed our experiments.

tral context, but subjects make systematic updating errors in the algebra/verbal test. The systematic mistakes tend to go against self-serving beliefs, as, unlike in Eil and Rao (2011), here people are more affected by bad news than by good, resulting in pessimistic beliefs. A difference from our study pertains to the type of feedback that is given. In her case, as in Eil and Rao (2011), feedback is always correct but incomplete (e.g., subjects are told they are not in the top 20%, but not if they are in the middle 60% or bottom 20%). Instead, we provide noisy feedback, which is not always correct. This makes it possible for participants to attribute any negative feedback to external causes, i.e., to 'bad luck' (noise).[7]

Burks, Carpenter, Goette and Rustichini (2010), based on data in Burks et al. (2009) investigate whether concerns for self-image contribute to overconfidence and whether confidence judgments are consistent with Bayesian information processing starting from a common prior, and reject both hypotheses. The results indicate that individuals with higher beliefs are more likely to demand information, rather than less likely. These results clearly reject self-image concerns as a mechanism that yields overconfident judgments, and are consistent with a model in which individuals like to hear good news. Burks et al. (2010) also derive restrictions that Bayesian updating places on the allocation function, and reject the hypothesis that confidence judgments are consistent with the Bayesian updating with a common prior.

Grossman and Owens (2010) study experimentally how one's beliefs about own performance (on a quiz) are affected by noisy, but unbiased feedback. In the main treatment, participants overestimate their own scores, believing that they have received unlucky feedback. However, this is driven not by biased information processing, but rather by overconfident priors. In a control treatment, each participant expresses beliefs about another participant's performance,

[7] Studies in psychology suggest that attributing success to one's self and attributing failures to external sources (such as 'noise') is a commonly employed strategy by people (see, e.g., Wolosin, Sherman, and Till, 1973)

with (on average) accurate posteriors. They also find that even though feedback improves estimates about the performance on which it is based, this does not lead to improved estimates of related performances. This result suggests that the manner in which people use performance feedback to update beliefs about own ability differs from the manner in which they update their beliefs about own performance, which may have bearing on the issue of why overconfidence persists.

## 3. Hypotheses and Experimental Design

The theoretical literature we have reviewed points to how incorrect favorable beliefs can be sustained, predicting that individuals can derive psychological benefits from overconfidence. The experimental literature suggests that in some cases people receive internal benefits from overconfidence, and pay more attention to good news than to bad news and wanting information only when they expect that they have done well; however, the Ertac (2008) and Burks et al. (2010) results call this into question. None of the existing studies has considered the relation of social and strategic considerations to overconfidence, and consequently none is able to make comparisons across the personal and social domains. This is what we do here. To provide the link between display of confidence in private and social situations we first state our theoretical hypotheses.

### 3.1 Hypotheses

The unifying view we adopt here is to consider the display of confidence as a special case of a display of emotional arousal. Expressions of confidence, like that of emotions, may be produced by the perception we have of own preparation for future actions, and of our own private information. This statement is a corollary of James' (1884) theory of emotions, according to which we first respond to an outside stimulus, and then we become aware of our own response. Emotions and their expression are, in this view, costly signals of our future intentions and private information. When a third party observes these emotions, he can make inferences about our future actions and of our private information. If we are in some measure aware of the effect of our emotional display on others, we can then adjust the emotion to the purpose described above. The way in which we make this adjustment depends in a crucial way on whether we are consciously aware of this effect or not. If we are, we are likely to adjust our behavior (in the case we are considering, our display of confidence), finely tuning it to the current situation, and to the precise incentives in the environment we are in. If we are not, the adjustment will be less specific.

The way we display our confidence may be different if our attitude to competitive situations is different. In light of the evidence that males and females respond differently to competitive environments (see e.g., Gneezy and Rustichini, 2004; Gneezy et al., 2003; Niederle and Vesterlund, 2007), we consider the possibility that confidence display is different across genders.

Based on these principles, we formulate the specific hypotheses below.

**H1**: Statements of confidence are social signals of intentions or private information, and individuals take them into account when they observe the self-evaluations of others. This is a first-order awareness of social implications of self-confidence. Individuals may also anticipate this effect and adjust this signal accordingly, a second order awareness. Thus in our experiment we predict that the stated confidence levels will be higher in a strategic environment, where the setting is explicit and the advantages clear.

**H2**: Higher stated confidence levels will tend in our experiment to discourage potential entrants from entering the tournament; so accurate reporting of confidence levels is not an optimal strategy, even taking into account the incentives for accurate reporting.

**H3**: If either preservation of self-esteem or ego-utility is present, then confidence will have a role of self-signal, likely to occur even when no one else is watching. Thus, this theory predicts that we will observe overconfidence even when one's confidence level is unknown to other participants, despite incentivizing people for correct beliefs.

**H4**: Males will exhibit higher stated confidence levels than females, even after controlling for performance. In line with previous research, we also predict that males are more likely to enter the tournament than females, controlling for confidence.

### 3.2 Experimental design

Sessions were conducted in Amsterdam in November 2009 and March 2010 with 16 to 28 participants depending on the number of subjects showing up for the experimental session. Instructions were displayed on a computer screen and read aloud. Participants were told that their decisions would remain anonymous to the other people present unless explicitly indicated otherwise, and that they would receive their earnings in an envelope from a person in a different room who could only see login numbers and could not match these numbers to names or faces. Participants were paid for one task chosen at random.

We ran a total of 17 sessions with a total of 368 subjects; seven of Treatment 1 (N = 144), three of Treatment 2 (N = 68) and seven of Treatment 3 (three with low outside option, N = 60, four with high, N = 96). Sessions lasted for 40 to 50 minutes, with an average payment of €14 (of which €7 was a show-up fee). Sessions ended with a questionnaire. Almost all participants (97 percent) were undergraduate students (average age 21.9 years, standard deviation 3.06; see Table 1 for details), with the majority studying economics or business; 43 percent of these subjects were female.

In every treatment, participants were randomly allocated to groups of four individuals. In each group, two players were randomly labeled as A players, and the other two were labeled as B players; each A player was randomly matched with one B player. All participants received the same 15 questions taken from Raven's Advanced Progressive Matrices (APM), a measure of cognitive ability (Raven, 2000). Participants had eight minutes to answer as many questions as they could, and did not get any feedback after completion on the number of questions they answered correctly. The timeline of the experimental sessions is shown in Appendix A, while the experimental instructions can be found in Appendix B. Payments were presented in points: One point was worth one euro. In the period in which the experiment was run 1 euro was worth approximately between 1.30 and 1.40 US dollars. In the exposition below we translate directly points into euros, although the instructions were strictly in terms of points.

First, when taking the APM test, participants only knew that they would be asked to evaluate their performance later and that every player A would be matched to a player B with a possibility for the player with the higher rank to earn 10 points, that is, 10 euros. Upon completion, participants were informed about all the subsequent steps in the experiment. First, one was asked to indicate one's confidence of having a score in the top two of their group, on a probability scale from 0% ("certainly not among the top 2") to 100% ("certainly among the top 2"). They received payment for accuracy according to a quadratic scoring rule; for a stated probability $p$ (their report divided by 100), a subject was paid 10 euros times $1 - (1 - p)^2$ if he really was in the top 2, and 10 euros times $1 - p^2$ if he was not. In the baseline treatment, no one could see the confidence of another player, each B player could observe the reported confidence by the paired A player in the other treatments.

Furthermore, in all treatments there was a possible tournament between the paired A and B players, in which the player with higher rank received 10 euros and the other received nothing. Entry by both players was mandatory, except in treatment 3 where each B faced a strategic decision: After observing A's reported confidence, B chose whether or not to enter a tournament, where the person with the higher rank received 10 euros, while the other received nothing. In the low-outside-option version of treatment 3, B received 3.5 euros by staying out, while in the high-outside-option version of treatment 3, B received 5.5 euros for doing so; in both cases, A received 10 euros for winning the tournament.[8]

The description we have just given, including whether or not any player could see the reported confidence of others, or whether player B was given a choice between playing in or out, was common information and known to all subjects before they reported their confidence. They were also told, in all treatments, that they would find out at the end of the game who had the higher rank between the two matched A and B players, but neither their rank in the group of four, nor the number of questions answered correctly.

Treatment 1 had some additional components, the Reports and the Machine questions, that we present now.

*Reports*: First, after reporting their confidence, participants were sent a report telling them if they were among the top 2 of their group or not. They were told that the report was always correct when it stated that one was among the top 2, but that the report was incorrect in 50% of the cases when one was in the bottom 2. There was no deception: reports to participants were determined as stated. After receiving the report, they were asked if they thought it was more likely that the report was correct or incorrect (i.e., if they were in the top 2 or not). They received 10 euros for a correct assessment.

*Machine questions:* Subsequently subjects were given an abstract scenario. Two machines in a production hall produce rings: the left machine produces 50% good rings and 50% bad rings, the right one produces 100% good rings. A mechanic inspects one of the machines every day by taking a ring. Participants were told the percentage of days the mechanic went to

[8] Treatment 2 is an intermediate treatment between 1 and 3: in this treatment, B cannot make a strategic decision (similar to Treatment 1), but player's A reported confidence is shown to B (as in Treatment 3). It is designed as a control treatment, to rule out that differences between Treatments 1 and 3 are not due to the visibility of A's reported confidence to B.

the left machine and were then asked if, given that the mechanic took a good ring, it was more likely that the mechanic went to the left machine or right machine. This question was asked three times, with various percentages of days that the mechanic went to the left machine. For the first two questions, the percentage was randomly drawn from a uniform distribution between 45 and 85. For the third question we used reported confidence levels by participants as an input. Participants were not told how the percentages were selected. One of the three questions was randomly selected, and for that question they received 10 euros for a correct assessment.

The reports and the abstract questions have an identical statistical structure (Figure C1 of Appendix C). In both, there are two possible states of nature, $\omega_1$ and $\omega_2$. The probability of state $\omega_1$ occurring is $p$. One of two signals was sent, $s_1$ or $s_2$. If $\omega = \omega_1$, each signal was sent with equal probability .5; if $\omega = \omega_2$, signal $s_2$ was always sent. Since $s_1$ only occurs when state is $\omega_1$, upon observing $s_1$, state $\omega_1$ must have occurred. Upon observing $s_2$, $\omega_2$ has posterior probability larger than $\omega_1$ if and only if $p < 2/3$.[9] In the report part, the states are $\omega_1$ = 'in top 2', $\omega_2$ = 'not in top 2', $p$ is the reported confidence, and $s_1$ = 'in top 2', $s_2$ = 'not in top 2'. In the machine part, $\omega_1$ = 'left machine', $\omega_2$ = 'right machine', $p$ is given in the question, and $s_1$ = 'bad ring', $s_2$ = 'good ring'.

# 4. Experimental Results

### 4.1 Confidence

Summary statistics are reported in Table 1. The distribution of correct answers (out of 15) is approximately normal, with mean 8.75 (8.77 for males and 8.73 for females). There are no

[9] If a participant received the report 'not in top 2', the likelihood that the report was incorrect (i.e., the person is in the top 2) is $.5p/(.5p + 1 – p)$ and the likelihood that the report was correct (i.e., subjects is not in top 2) is $(1 – p)/(.5p + 1 – p)$. Conditional upon receiving the report 'not in top 2', incorrect (top 2) is thus more likely than correct if and only if $.5p > 1 – p$, that is $p > 2/3$. Similar reasoning shows that upon receiving the information that the ring is good, it is more likely that the mechanic went to the left machine than to the right machine if and only if $p > 2/3$, where $p$ in this case is the probability that the mechanics went to the left machine.

large or significant differences of reported confidence levels across conditions: the pairwise *p*-values are in all cases larger than 10 per cent.[10] In all conditions less than 1/3 of the participants report a confidence level of 50 percent or below.

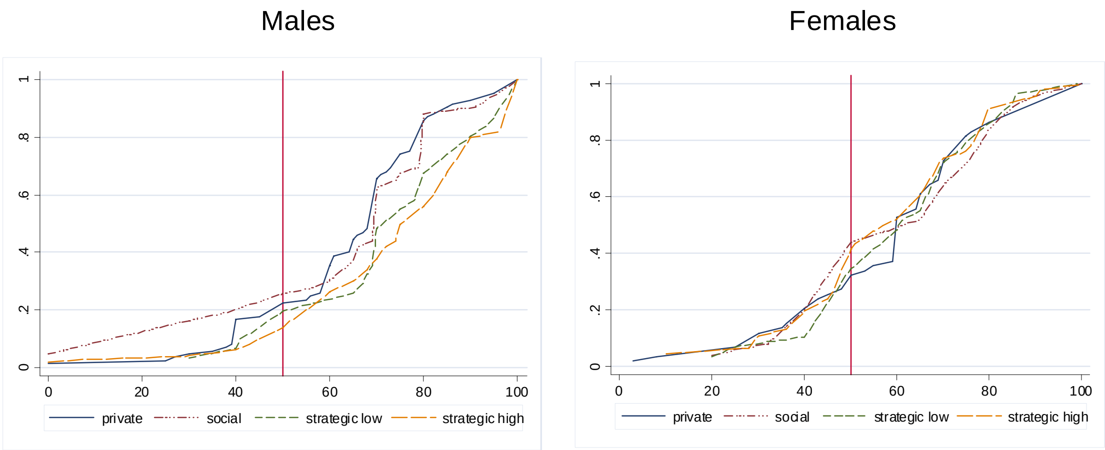
**Table 1: Summary statistics**

| | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|
| **Test** | | | | |
| Number correct answers | 8.75 | 2.44 | 3 | 15 |
| Confidence | 65.23 | 21.14 | 0 | 100 |
| | | | | |
| **Background characteristics** | | | | |
| Age | 21.90 | 3.07 | 17 | 49 |
| Number siblings | 1.47 | 0.99 | 0 | 6 |
| Gender (fraction females) | 43% | | | |
| Member of sports club | 52% | | | |
| Took Raven test before | 52% | | | |
| Familiar with conditional probs. | 60% | | | |
| | | | | |
| **Study category** | | | | |
| Economics/Business/Finance | 59.56% | | | |
| Social Sciences and Law | 15.30% | | | |
| Physics, Math, Computer science | 6.56% | | | |
| Other study or not student | 18.58% | | | |
| | | | | |
| N | 368 | | | |

Figure 1 reports the cumulative distribution of reported confidence by condition and gender. It appears that males report higher confidence in the strategic treatments (left panel), while there are no discernible treatment effects for females (right panel).

[10] They are: 0.717 (Treatment 1 and 2), 0.300 (1 and 3-Low outside option), 0.136 (1 and 3-High), 0.538 (2 and 3-Low), 0.362 (2 and 3-High), and 0.994 (3-Low and 3-High).

**Figure 1: Cumulative distribution functions of confidence, by treatment.**

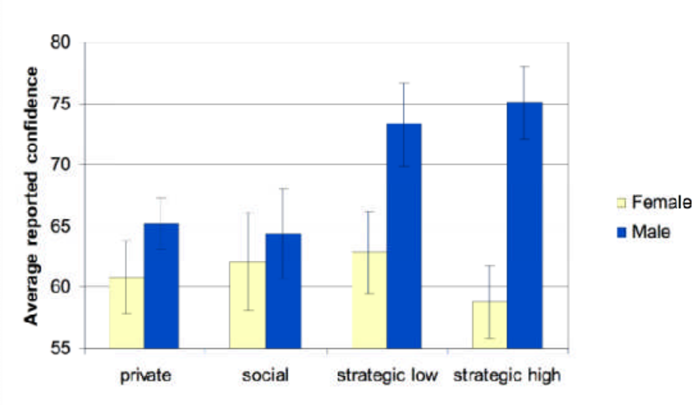



"Private" refers to Treatment 1, "social" to Treatment 2, "strategic high" to Treatment 3 with the high outside option of 5.5, and "strategic low" to Treatment 3 with the low outside option of 3.5. Left panel: males; right panel: females. The vertical bar is at 50 per cent.

In data pooled over the conditions, 72 percent of the people report a confidence level above 50 percent. Figure 2 shows a treatment and gender effect with respect to mean reported confidence. Men are considerably more confident in the strategic condition compared to the other treatments ($p$ = 0.001, WMW test), while no difference is found for women ($p$ = 0.927). Note that the participants are told about the strategic interaction and asked for their confidence *after* taking the Raven test, so that performance *per se* could not be affected by awareness that a strategic decision would follow, whereas statements might be affected.

OLS estimates of the determinants of confidence are presented in Table 2; the baseline condition reflects male behavior in the private treatment (Treatment 1). Specification (1) shows that the number of correct answers is a strong predictor of confidence, adding about 3 percentage points for each correct answer; this result is robust over different specifications.

**Figure 2: Confidence of being in top 2, by condition and gender.**




See Figure 1 for descriptions of the variables.

Since subjects were not told their number of correct answers, the effect of correct answers on stated confidence is not based on any external information, but on an estimate of one's own ability positively related to real ability. Specification (2) adds controls for the role of the subject (player A or B) and interaction terms for the treatment and role. Being a player A has no effect by itself, and there is also no significant interaction effect with the treatments.

Introducing a dummy for gender and interaction effects for gender and treatment in specification (3) shows a negative but insignificant gender effect. There is a significant treatment effect: Treatment 3 increases reported confidence by almost 10 percentage points. This effect is only present for males: the coefficient of the interaction between Treatment strategic and female shows a negative coefficient of about equal size as the treatment coefficient. There is no difference for player A or with the interaction of treatment 3 and player A, and no difference for the three-way interaction female, player A and treatment 3. Specification (4) shows that people who indicated they were familiar with conditional probabilities are more confident, by more than 4 percentage points. This familiarity has a significant effect even accounting for the difference in the number of correct answers (9.31 versus 8.14).

**Table 2: Determinants of confidence (0 – 100) -- OLS estimates**

| Dependent variable: Confidence | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Number of correct answers | 3.26*** | 3.27*** | 3.29*** | 3.10*** |
| | (0.42) | (0.42) | (0.41) | (0.43) |
| Social | 0.28 | 0.59 | 0.62 | -1.49 |
| | (2.88) | (4.10) | (3.99) | (4.02) |
| Strategic (low and high option combined) | 3.59 | 3.86 | 9.98** | 9.70** |
| | (2.26) | (3.21) | (3.94) | (3.95) |
| Player A | | 1.65 | 2.06 | 2.12 |
| | | (3.28) | (3.20) | (3.16) |
| Treatment social * Player A | | -0.63 | -0.96 | 0.65 |
| | | (5.82) | (5.66) | (5.68) |
| Treatment strategic * Player A | | -0.54 | -1.60 | -1.61 |
| | | (4.54) | (5.32) | (5.32) |
| Female | | | -3.23 | -2.76 |
| | | | (2.70) | (2.70) |
| Treatment strategic * Female | | | -10.88** | -12.10** |
| | | | (5.09) | (5.09) |
| Treatment strategic * Player A * Female | | | -0.56 | -1.05 |
| | | | (6.14) | (6.13) |
| Familiar with conditional probabilities | | | | 5.43*** |
| | | | | (2.10) |
| Constant | 35.10*** | 34.18*** | 35.19*** | 29.99*** |
| | (3.98) | (4.42) | (4.43) | (9.35) |
| Controls | No | No | No | Yes |
| N | 368 | 368 | 368 | 366 |
| Adj. $R^2$ | .14 | .14 | .19 | . 22 |

Control variables: familiarity with Raven test, study category, age, number of siblings, birth order, member of sports club, entity theory question. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

B players inflate their confidence levels to the same extent as the A's. Their inflated levels of stated confidence cannot deter A players from competing, but may reflect strategic motivations generated by the competitive setting when people are not flexible enough adjust their behavior to the contingent role (A or B) in the strategic environment. The notion of competition can trigger a reaction, without consideration for how payoffs are determined in the specific situation, a phenomenon that has been previously documented.[11]

*Results (confidence):*

1. *The real performance of participants, measured by the (unknown to the participants) number of correct answers, greatly influences reported confidence in the expected direction. Those people who are familiar with conditional probabilities also report higher confidence, after controlling for correct answers.*
2. *There is a significant treatment effect for men, who report 10 percentage points higher confidence in Treatment 3 (where confidence is observed and may have strategic effects), even though it is only known after taking the test that there will be strategic interaction (but before the statement is given). There is no significant treatment effect for women.*
3. *Player A does not report higher confidence than Player B does in any treatment, whether for males or females. In Treatment 3, the similarity of the behavior of players in the two roles may reflect automatic response to competition on an unconscious level.*

The confidence reports in Treatment 3 will be discussed again in the analysis of the strategic behavior of participants.

[11] Fershtman, Gneezy, and List (2009) provide experimental evidence that the mere presence of a competitive environment can trigger strong behavioral changes. While people exhibit typical social preferences in a standard dictator-type decision, the same people who were generous in this baseline behaved very differently when relative performance on a task determined the allocation, as they competed hard for the more remunerative (and less egalitarian) outcome. Similarly, in Liberman, Samuels, and Ross (2004) participants play a Prisoner's Dilemma game that is labeled either the "Wall St. Game" or the "Community Game". People were significantly more likely to choose defection when they played the Wall St. Game. They react to the labels that are associated with different degrees of competition, even when the monetary payoffs of the underlying game are unchanged.

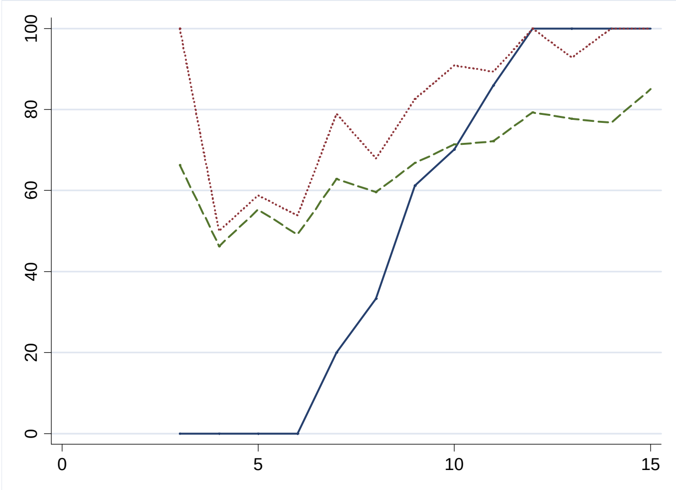
### 4.2 Overconfidence

The purpose of the paper is to provide explanations for an empirical fact, that the fraction of people claiming to have a score higher than a given percentile is larger than the fraction of those who are in reality above that threshold. We call this fact overconfidence, and we do not imply a bias with respect to the Bayesian benchmark. The existence of such a bias is a possibility we test, by checking whether the discrepancy can occur with Bayesian updating. Within our data we cannot reject the possibility that the reported confidence levels are consistent with Bayesian updating. A simple test of overconfidence is to compare the fraction of subjects that report confidence above 50 to the fraction of subjects that report confidence below 50. With a binomial test, we can easily reject the null hypothesis that these proportions are equal ($Z$ = 6.41, 4.30, 7.89, and 10.98 for Treatments 1, 2, and 3 and all treatments pooled, respectively; each of these test statistics are significant at $p$ = 0.000.). Similarly, the fraction of people with confidence above 50 is significantly higher than 50 percent in all cases.[12]

Figure 3 shows the proportion of subjects actually in the top two of their group by the number of correct answers (solid line). No one with fewer than seven correct answers is in the top two, while everyone with more than 11 correct answers is in the top two. The dashed line shows average confidence according to the number of correct answers. Those with few correct answers appear overconfident, while those with many correct answers appear under-confident. The dotted line shows the percentage of people with reported confidence above 50; the percentage of people who think they are most likely in the top 2 always exceeds 50%.

[12] We can also compare confidence estimates with actual performance. Of all participants who are in the top two of their 4-person group, 86 percent report confidence above 50, versus 72 percent of those who are not in the top two. Thus, we cannot reject rational Bayesian updating using the test of Burks *et alii* (2009), so that we cannot conclude that subjects deviate from a common prior, Bayesian updating and truthful statements of their estimated relative position. Our interest in this paper is however in the difference in reported confidence level between treatments, not in the relation between real performance and stated confidence.

**Figure 3: Number of correct answers and confidence**




Number of correct answers in the Raven test on the horizontal axis. On the vertical axis: Actual percentage of subjects in top 2 (solid line), average reported confidence (dashed line), percentage of subjects with confidence above 50 **(**dotted line).

*Results (overconfidence):*

1. *The average reported confidence is 65 (and the median 69) rather than 50.*
2. *The average reported confidence is systematically higher than the actual percentage in the top 2 for participants with relatively few correct answers, and it is lower for participants with relatively many correct answers.*

We cannot reject rational Bayesian updating using the Burks et alii (2009) allocation function. This may reflect our having only two intervals, either above or below the median.

### 4.3 Updating errors

Comparing choices across the Reports and Machine questions tasks in Treatment 1 allows us to test the hypothesis that the patterns of stated confidence that we observe are due to errors in Bayesian updating. Recall that each person received a report about his or her rank in the group, and then answered three machine questions, each with a different percentage of days that the mechanic went to the left machine.[13] With this information, we can test whether mistakes in the Reports task are due to a general inability to update posterior probabilities.

Since a report that one is in the top 2 can only occur when one really was in the top 2, such a report is fully informative, should be believed, and it always is in the data. So we focus on the response to negative feedback ("not in the top 2"). There is a threshold prior confidence level of 2/3 for whether or not one should believe one is in the bottom 2 after a negative report. Thus, an updating error for someone with initial confidence below this level would consist of thinking that one is more likely to be in the top 2 after receiving negative feedback; for someone with initial confidence above the threshold, the error would consist of believing the opposite. The two errors represent respectively a form of under-updating and over-updating.

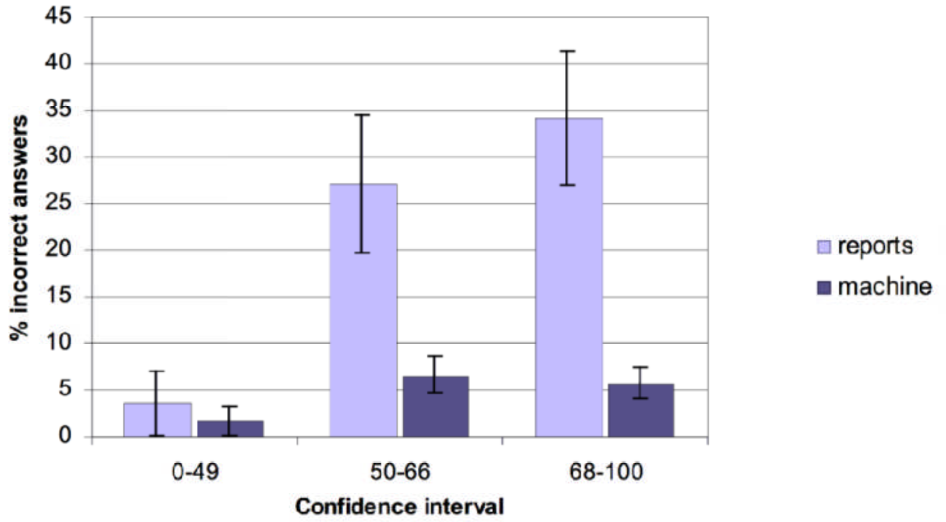
The test that errors are due to a general inability to update is provided by the comparison between the frequency of errors when belief is about one's relative skill and when it is about an abstract evaluation. We find evidence of significantly more frequent errors in the Reports than in the Machine questions. Figure 4 compares mistakes made by subjects in each of those intervals after they receive negative feedback in the reports, compared to the mistakes with the machine question.

There are very few errors for people with stated initial confidence below 50, regardless of the question. Updating mistakes are always rare with the Machine questions, as the percentage of mistakes ranges from less than two percent to less than seven percent. The picture is very different for the Report questions: 27 percent of subjects make under-updating errors in the interval between 50 and 67, and 34 percent of subjects believe negative feedback when stated confidence

[13] Note that no feedback about actual performance is given, so that nothing is learned and the estimates given for the Machine questions seem independent of the estimate given for the Reports question. There is no obvious reason why the order of these choices should matter; if anything, one might expect the estimate for the Machine question with the same probability of visits as the level of confidence reported to be sticky, rather than changing.

was larger than 67 percent.[14] Statistical tests confirm these error rates on the Report questions are significantly higher than with the Machine questions ($Z$ = 3.72 and Z = 5.46 for the respective comparisons, both $p$ = 0.000). Furthermore, if we restrict the sample of machine questions to the third round, where reported confidence levels are used as an input, the difference in error rates between the reports and machines questions are still significant at the 5% level ($Z$ = 1.78, $p$ = 0.038 and $Z$ = 3.87, $p$ = 0.000, both one-tailed tests).

**Figure 4: Updating errors rates for Reports and Machine questions, by confidence interval.**
An updating error is defined as an answer differing from the one provided by Bayesian updating of a prior equal to the stated confidence of the subject




What is the source of errors in the Reports? It is not the case that our participants are unable to do the updating task. Their performance on the machine questions shows otherwise. Participants making updating errors on the Reports seem to do worse in terms of performance. If we group over-updating and under-updating errors, we find that those making errors have fewer cor-

[14] The fractions of errors in the intervals 50-67 and above 67 are not significantly different from each other ($Z$ = 0.69, $p$ = 0.493).

rect answers on the Raven test: 7.6 is the mean among participants making errors versus 8.3 of those who do not. This difference is not significant (the WMW test gives $Z = 1.10$, $p = 0.272$). However, once we control for reported confidence, the difference becomes significant (column 1 in Table 3).[15] Thus, a lower cognitive ability is associated with making updating errors. We find a similar effect for errors on the Machine questions but the coefficient is not significant (column 2). In the case of the Reports, the worse performance translates into a large effect on the likelihood of actually being in the top 2 (19 percent versus 41 percent, $Z = 2.00$, $p = 0.045$), but not in the case of errors on the Machine questions 2 (32 percent versus 37 percent, $Z = 0.41$, $p = 0.680$).

**Table 3: Determinants Updating errors and performance on the Raven task (over-updating and under-updating errors combined)**

| Dependent variable | Number of correct answers on the Raven test | |
|---|---|---|
| | (1) | (2) |
| Report error | -1.09** | |
| | (0.522) | |
| Machine error | | -0.90 |
| | | (0.586) |
| Confidence | 0.04*** | 0.03*** |
| | (0.011) | (0.011) |
| Constant | 6.09*** | 6.49*** |
| | (0.682) | (0.721) |
| N | 112 | 112 |
| R-squared | 0.11 | 0.10 |

Sample is participants in Treatment 1 that received negative feedback.
Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

[15] If we estimate model (1) of Table 3 separately for both types of errors, the coefficients remain similar but become insignificant on the intervals 50-67 and above 67, possibly due to a lack of power. Nonparametric tests show that participants making over-updating errors (interval above 67) do significantly worse on the Raven test ($Z = 1.776$, $p = 0.076$), but there is no significant effect for under-updating errors (interval 50-67, $Z = 1.107$, $p = 0.269$).

*Results (updating errors):*

1. *Errors in updating are much more likely to occur for questions involving own skill judgment than for the abstract questions; hence such errors do not merely reflect a general inability to perform Bayesian updating. Errors in response to the abstract questions are rare.*
2. *There is no significant difference in the rates of under-updating errors and over-updating errors.*
3. *People who make errors on the Reports do worse on the Raven test and are much less likely to be in the top 2, controlling for confidence.*

**4.4 Voluntary tournament entry**

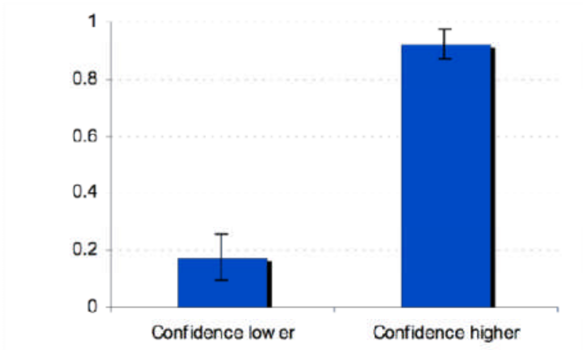
In Treatment 3, player B chooses whether to enter a tournament with player A, who is automatically entered into the tournament. This result contrasts with Treatments 1 and 2, where both people are effectively entered into a tournament. B decided whether to enter after he had seen A's reported confidence, so he could take that confidence into account when estimating the probability of winning the tournament. Player A in turn knew that player B would observe his statement at the moment of stating his confidence (and player B knew this, etc., since the instructions were identical for all participants), and could anticipate the effect of the statement on player B's decision. A does not observe B's statement, so this statement could not affect A's behavior. In light of this, what determines player B's choice?

Our data show that with the high outside option, player B is more likely to enter the tournament when own confidence is higher and when the opponent's confidence is lower.[16] Indeed, as is shown in Figure 5, we observe that relative confidence is a phenomenally good predictor of entry. Twenty-three of 25 B's (92.0%) enter when their confidence level is at least as large as

[16] We focus primarily on entry with the high outside option, since 28 of 30 B's chose entry with the low outside option, so that statistical tests have little power.

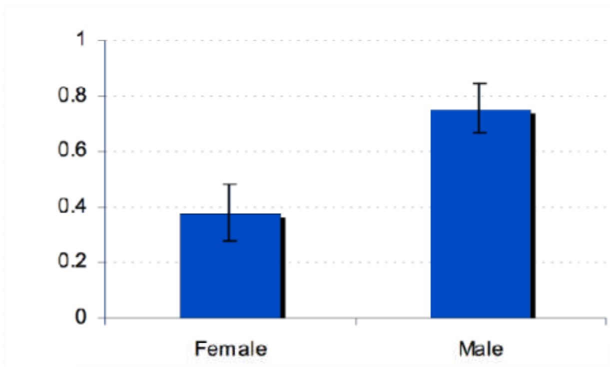
the paired A's reported confidence level, while only four of 23 B's (17.4%) enter when their confidence level is lower than the paired A's reported confidence level; the difference in these proportions is highly significant ($Z$ = 5.21, $p$ = 0.000). Thus, there is strong potential for strategic deterrence on the part of the A player.

**Figure 5: Entry rates by higher/lower confidence with high outside option.**




We also find that males enter twice as frequently as do females, 75.0 percent versus 37.5 percent ($Z$ = 2.62, $p$ = 0.009, two-tailed test), as is shown in Figure 6. However, this does not reflect a difference in performance: females in the B role in the high-option condition do nearly as well as males on the Raven test (the mean score for males is 9.12 and the mean score for females is 8.88; WMWtest: $Z$ = 0.28, $p$ = 0.779, two-tailed test).

**Figure 6: Entry rates by gender with high outside option.**



At first glance, this seems to be evidence that females are *per se* averse to competition. However, female B's state significantly lower confidence levels than do male B's in this condition, 58.76 versus 75.04 ($Z = 3.07$, $p = 0.002$, two-tailed test). Men choose to enter a tournament twice as often as women do, but this reflects a lower stated confidence level. This effect is only seen for people who choose to enter the tournament; the average stated confidence level for male entrants is 84.17 versus 69.89 for female entrants, while the average stated confidence level for male non-entrants is 50.83 versus 48.67 for female non-entrants.

Table 4 reports the probit estimates of the decisions to enter the tournament for the high-option sessions of Treatment 3.[17] Specification (1) shows that own confidence increases the likelihood of entering the tournament, while higher reported confidence by the opponent lowers the likelihood of entering. Each variable substantially affects the probability of entering. Specification (2) includes a dummy variable that simply compares if own confidence is higher or lower than that of the opponent. This is a very good predictor.

The next two specifications add some controls. From (3), we see that the gender coefficient is negative but not significant. Finally, controls in (4) for gender, number of correct answers, and risk aversion have no significant effect. Thus, the fact that females are less likely to enter is mainly driven by the fact that they are less confident, rather than less competitive. This also suggests that males are not just reporting higher confidence, but also feel more confident. If they were just reporting higher confidence without believing it, then, controlling for confidence, males should have been less likely to enter the tournament.

[17] Estimates from the Linear Probability Model are very similar to the reported Probit marginal effects.

**Table 4: Determinants of entering tournament, Probit estimates, marginal effects**

| Dependent variable | Choice = IN | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Sample | High outside option | High outside option | High outside option | High outside option |
| Own confidence | .044*** | .015** | .015** | .014* |
| | (.012) | (.007) | (.007) | (.008) |
| Opponent's confidence | -.029*** | | | |
| | (.010) | | | |
| Lower confidence | | -.619*** | -.610*** | -.621*** |
| | | (.138) | (.141) | (.141) |
| Female | | | -.121 | -.096 |
| | | | (.199) | (.203) |
| Number correct answers | | | | -.019 |
| | | | | (.042) |
| Risk aversion† | | | | -.029 |
| | | | | (.080) |
| N | 48 | 48 | 48 | 48 |
| Pseudo $R^2$ | 0.59 | 0.56 | 0.57 | 0.57 |

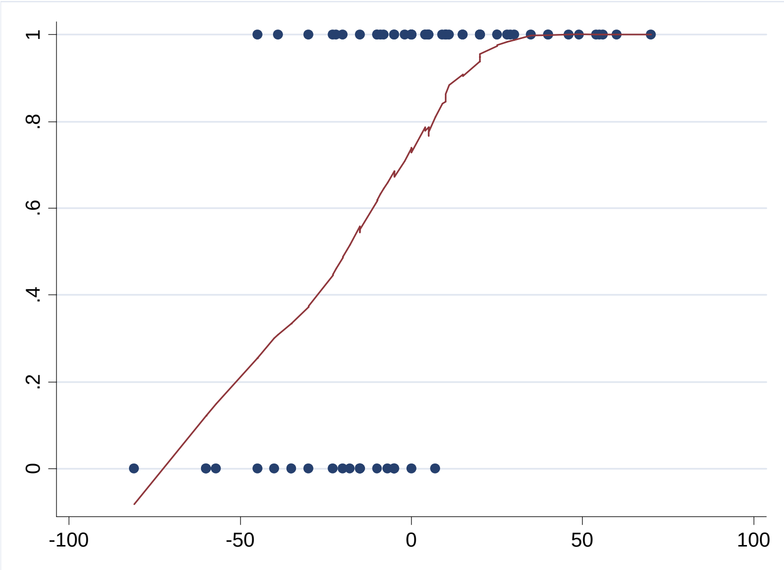
†Eight missing observations were replaced by the mean. Lower confidence is equal to 1 if B's confidence is lower than the reported confidence of the paired A, and is 0 otherwise. Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

*Results (tournament entry):*

1. *When deciding whether to enter the tournament, participants are more likely to enter when their confidence is higher; they are also sensitive to the confidence reported by the opponent: If own stated confidence is lower than that of the opponent, subjects are far less likely to enter.*
2. *Females are less likely to choose IN and enter the competition, but this effect is mainly due to the difference in confidence. Once we control for confidence, the entry rate of women is not significantly lower.*

Finally, Figure 7 illustrates graphically a Lowess-smoothed estimation of B's reaction function to the difference in own and other reported confidence. We will use this response function of B when analyzing the optimal decisions for player A in the next section.

**Figure 7: Estimated rate of B entry by difference in reported confidence**




On the horizontal axis is the difference between B's own reported confidence $(r_B)$ and reported confidence by A $(r_A)$, i.e., $r_B - r_A$. The curve shows the Lowess smoother estimate with bandwith 0.8.

## 5. Discussion

In this section, we discuss the extent to which our data indicate the presence of self-esteem, social concerns, and self-deception. We then proceed to an analysis of the A player's optimal behavior in the high-outside-option condition, given the B player's anticipated responses. Finally, we discuss whether strategic overconfidence can be an equilibrium in our setting. Indeed, we find that strategic overconfidence can survive in equilibrium.

### 5.1 Overconfidence and its motivations

One key potential motivation for being overconfident is the ego utility that one derives, corresponding to an increase in self-esteem. Our data provide some support for this notion in the expressed level of confidence, but not in the updating. We observe substantial overconfidence even when the stated confidence level is not observed by anyone else, suggesting that people are either poor judges of probabilities or that they receive some internal benefit from this inflated belief. The fact that people make far fewer updating errors with the Machine questions than with the Reports questions suggests that the explanation of the overconfidence is not poor ability to estimate probability of events. On the other hand, errors in updating do not exhibit any bias in the direction suggested by ego-utility or preservation of self esteem. Many subjects with a prior stated probability lower than the threshold update and receive a negative report still state a probability larger than 50% that they are in the top 2, and this might be considered a bias induced by self-mage preservation. But correspondingly many subjects who should, according to their maintain a probability larger than 50% do not do so facing a negative report, a bias that goes contrary to self-image preservation.

Regarding social concerns there is strong evidence that an increase in A's reported confidence can have deterrence value in terms of keeping B out of the tournament. We also see some evidence (see Figure 1and Table 2) that males report higher confidence in the strategic condition than in Treatments 1 and 2. How close is this behavior to that which is optimal for the A players? We take this up in the next subsection.

### 5.2 Determining optimal confidence reporting in the strategic setting

Since we have already established that B players are sensitive to reported confidence by A players, we investigate the best response of player A. We take the response of players B in the game (as shown above in Figure 7) and compute the best response of player A, under the assumption that he knew this response function. We then compare the best response of player A to the actual behavior.

To analyze the best response of player A, denote by $t_A \in [0,100]$ and $r_A \in [0,100]$ the true confidence and the reported confidence respectively by player A, and similarly $t_B$ and $r_B$ for B. The distribution of $t_B$ is assumed to be normal with density $f(t_B)$ and c.d.f. $F(t_B)$ and truncation points 0 and 100. Let $P_B(IN)$ be the probability that player B plays IN. This depends on the reported confidence of A. Based on Figure 7, we estimate player B's response function as:

$$P_B(IN) = \begin{cases} 1-\beta(r_A - t_B) & if\ r_A > t_B \\ 1 & if\ r_A \leq t_B \end{cases}$$

where we estimated $\beta = .01$.

If $\Pi(t_A,t_B)$ is the probability that player A wins if B plays IN, the expected payoff for player A of the tournament is given by:

$$\int_0^{100} 10[1 - P_B(IN)\cdot(1-\Pi(t_{A,}t_B))]dF(t_B)\,.$$

Player A can expect to earn 10 euros, except if B plays IN *and* wins. The benefit of a higher report is the decreased probability that B plays IN. On the other hand, since players are incentivized by the quadratic scoring rule, there are costs associated with reporting a higher confidence than the true confidence. The optimal reported confidence for a risk-neutral player A is determined by the following equality:

$$r_A^* = t_A + \tfrac{1}{2} c \int_0^{r_A^*} \beta (1 - \Pi(t_{A,} t_B)) dF(t_B) .$$

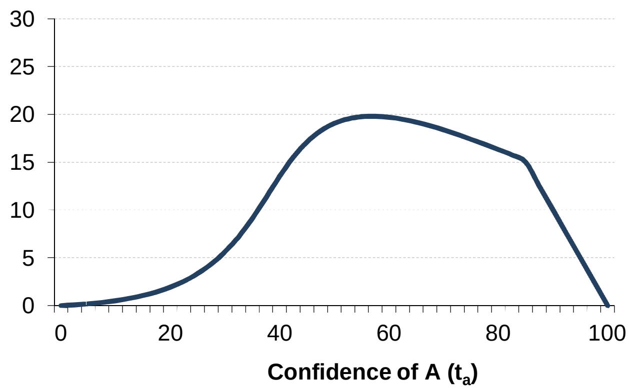
The constant $c$ = 10,000. This, and the '1/2' come from the quadratic scoring rule. For our estimation of $r_A^*$, we specify $\Pi(t_A, t_B)$ as a logistic function:

$$\Pi(t_A, t_B) = \frac{e^{\delta(t_A - t_B)}}{1 + e^{\delta(t_A - t_B)}}$$

For equal abilities, each player has a probability .5 of winning. From the Treatment 1 data (with no incentives to over-report) we obtain an estimate of $\delta$ = .025.

Finally, we need to make assumptions about the density function $f(t_B)$. We assume that true confidence is normally distributed (truncated at 0 and 100) with mean $\mu$ and standard deviation $\sigma$. In treatment 1 there is no scope to over-report for strategic reasons for any of the players. In this treatment, the mean and the standard deviation are given by $\mu$ = 65 and $\sigma$ = 21. However, as a benchmark we will assume that players believe the mean to be $\mu$ = 50 so that they don't believe that other players are on average overconfident.[18]

**Figure 8: Optimal over-reporting of confidence in the high-option strategic setting.** On the vertical axis we report the difference between the optimal report and the real confidence.




[18] Results are essentially unchanged assuming a true mean of 65.

We use the above estimates as the benchmark case. Using these parameter estimates, we can derive the optimal overconfidence $r_A^* - t_A$, which is shown in Figure 8. The curve is largely unchanged with different parameter values. We see that player A has an incentive to over-report by an amount that 1) is low or zero for small values of the true confidence; and 2) has a peak at intermediate values, where the report is as much as 20 percentage points higher than the true confidence level. For participants with true confidence above 85, the optimal report is 100, so optimal over-reporting declines linearly to zero as the confidence increases in this range.

Is the amount of over-reporting (in the strategic setting versus the baseline treatment) comparable to the best response we calculated? According to Table 2, male A players report approximately 8.4 percentage points higher confidence in the high-option strategic setting, while the confidence reports for female A players do not differ across the strategic and non-strategic settings. The fact that reported confidence of male players A is sensitive to whether or not player B can respond to this signal, suggests that confidence is indeed used as a social signal to influence others. Given the actual distribution of reported confidence in Treatment 1, we calculate that the average optimal degree of over-reporting relative to Treatment 1 is about 14.1 percentage points.

### 5.3 Determining the optimal entrance decision

It also appears that B players may realize that A players over-report, and deflate these statements. In specification (1) in Table 4, the coefficient 'own confidence' is almost 50 percent higher than 'confidence other' in absolute terms, so one's own feeling is more important for the decision to enter than that of the opponent; this difference is significant at the 10 percent level.

To examine if B players are best responding, assume they believe that players A over-report in accordance with estimates from the previous section (see Figure 8). Let $\mu(t_A | r_A)$ be the beliefs that B assigns to a type of A ($t_A$) after observing $r_A$. With an outside option of ω, a risk neutral player B is indifferent between IN and OUT if $t_B = \tilde{t}_B$, where (provided $\tilde{t}_B \in [0,100]$):

$$\int_0^{100} 10(1-\Pi(t_A,\tilde{t}_B))\mu(t_A|r_A)dt_A \equiv \omega.$$

Any player B with $t_B < \tilde{t}_B$ will opt out. If all beliefs are concentrated on a single type $t_A$, then:

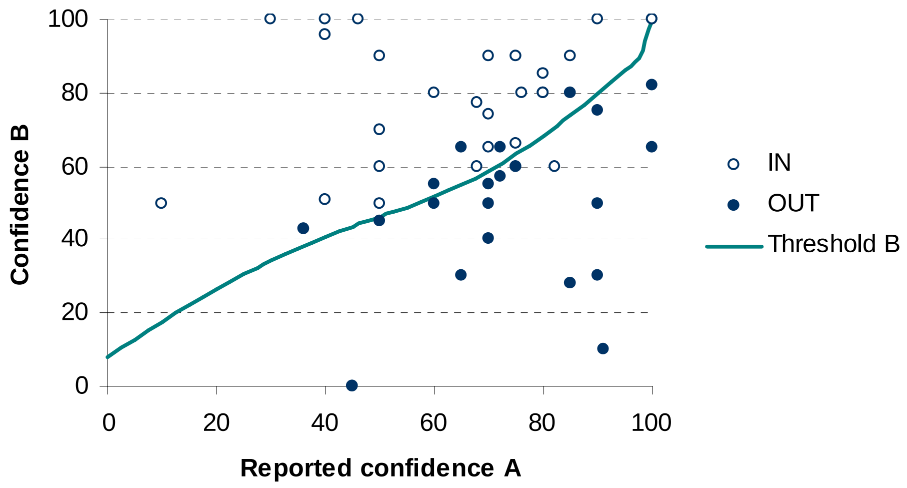
$$\tilde{t}_B = t_A - \frac{1}{\delta}\ln\left(\frac{10-\omega}{\omega}\right).$$

If ω = 5.5 (high outside option) and δ = .025, then $\tilde{t}_B \approx t_A + 8$. If several types of A players pool and report the same confidence, then B computes the expected payoff based on the conditional distribution of all types in the pool.

Let the inferences that B makes after observing A's report, i.e., $\mu(t_A|r_A)$, be based on $r_A^* - t_A$ as derived in the previous section. All types $t_A$ exceeding 84.7 will pool at a report of 100. For types below 84.7, the report is below 100 and the report reveals the true type of player A. Again, we will assume that players believe the mean confidence of others to be $\mu$ = 50 and a standard deviation of 21, so that they don't believe that other players are on average overconfident.

Figure 9 shows the threshold of B and IN/OUT decisions for the high-outside-option treatment. The reported confidence of A players is on the horizontal axis, and that of the B players is on the vertical axis. The dots represent all choices of the players. The players that chose OUT are represented by the solid dots, the IN choices by the open dots. The solid curve plots the estimated threshold level for B. If this were the true threshold level for the best response, all B players with a confidence above the threshold level should choose IN. The computed threshold level has strong predictive power, being consistent with 87.5% of the actual choices; furthermore, the few inconsistent choices are quite close to the threshold.

**Figure 9: Entry choices of players B in the high-option strategic setting**




### 5.4 Equilibrium

This takes us to the issue of whether strategic overconfidence is sustainable in equilibrium. In Figure 8, players A over-report, and the only pooling interval occurs for players with a confidence above 84.7, all reporting 100, while all others separate. Hence, players B can identify the true type from the reported type of A for reports below 100. Can it be that players A over-report in equilibrium even though players B can infer the true type of A, and even though over-reporting is costly for A? The answer is yes. The reason is that if players B believe that there will be over-reporting by A, then if some player A reports her true type, B will associate A with an even lower type of A player. This is analogous to games of limit pricing under incomplete information, where incumbent firms try to signal their strength to potential entrants by lowering their price. Lowering the price is a costly action, but may deter competition by making it appear unattractive to enter (see, for instance, Milgrom and Roberts, 1982). Such games can have equilibria with (partial) separation of types.

Consider, for instance, the following set of strategies.[19] It is natural to focus on a monotone equilibrium in which higher types of A report a (weakly) higher confidence. Let the 'strategy' of player B be to opt out if and only if $t_B < \tilde{t}_B$, where the threshold type $\tilde{t}_B$ is determined as above (as the type that is indifferent between entering or not given any observed confidence level of A and the corresponding beliefs about the conditional type distribution of A), and we assume that it has an interior solution. This threshold is increasing in the reported confidence of A under the monotonicity assumption above. For player A, let there be a threshold value $\tilde{t}_A$ such that all types $t_A \geq \tilde{t}_A$ pool and report the maximum confidence of 100, and all types below the threshold type separate. All types separating report a confidence level that satisfies the following condition:

$$\tfrac{1}{2}c[1-\Pi(t_A,\tilde{t}_B(r_A))]f(\tilde{t}_B(r_A))\frac{d\tilde{t}_B(r_A)}{dr_A} = r_A - t_A,$$

which is the first order condition of player A's payoff function with respect to the reported confidence. The threshold type $\tilde{t}_A$ is determined by the condition that this type must be indifferent between revealing herself (by following the separating strategy) and pooling with all higher types. This set of strategies may constitute a Perfect Bayesian Equilibrium if player B has sufficiently pessimistic beliefs about A's type after observing an out of equilibrium reported confidence. In such an equilibrium, all types are (weakly) over-reporting, i.e. reporting a higher type than their true type. For all types pooling at 100 this is evident. For the types separating, this can be seen from the fact that, to satisfy the above first order condition, $r_A$ must exceed $t_A$. Whether

---

[19] What follows is a sketch of an equilibrium with partial separation. Our game has many similarities to Kartik (2009), for instance, who analyzes a model of strategic communication with lying costs with a continuum of types.

or not this particular equilibrium exists depends on the specific parameter values, and, as is usually the case within the class of signaling games, many other perfect Bayesian equilibria may coexist.

## 6. Conclusion

Our experiments examined the determinants of self-confidence, and the degree to which it reflects self-image (consumption or ego utility), social image (utility from the perceptions of others, or strategic concerns. Our conclusion is that overconfidence is influenced by strategic interest, perhaps unconsciously processed.

Several results support this conclusion. The preponderance of the participants states a belief that they are above the median in their ability on the task. This is not consequence of a generic error in Bayesian updating because people are much more accurate in updating their beliefs when the information is about an impersonal issue then they are when new information concerns their own ability. In fact, they are hardly ever wrong in the abstract setting. This shows that information processing when own reputation (either in one's own eyes or the eyes of others) is at stake is of a different nature than abstract, neutral information processing. Why does this occur, and does it induce a bias in updating?

Our novel strategic environment (in which another party observes the stated confidence level of another and then chooses whether or not to enter a tournament with this other person) allows a direct test of the strategic concern hypothesis. First, the social signal is perceived and has consequences: subjects in our experiment do respond to statements about confidence made by others, taking that information into account when choosing whether or not to enter. This is particularly true for male participants, who in the 'incumbent' position on average report significantly higher confidence levels in this treatment than in the non-strategic ones. For males the observed level of (costly) inflated stated confidence is not too far from the estimated optimal level, and also serves as an effective deterrent. There is no such inflation of confidence observed for women. Our simple estimation of the threshold own-confidence value for entry does an excellent job of predicting actual entry decisions; relative stated confidence is a strong predictor of entry decisions. There is an equilibrium in which stated confidence is inflated (at a direct cost to the sender).

The alternative hypothesis that people err in the direction of overconfidence for self-image purposes does not find support in the pattern of errors they make. We do not find evidence of overall bias in the direction of positive self-image preservation, as the two types of updating errors are nearly equally common.

If strategic behavior in competitive environments affects confidence, it is natural to find gender differences in our subjects' behavior, since there is considerable evidence on gender difference and competition. In the private (non-competitive) environment, males state confidence levels only slightly higher than those stated by females, and have identical ability levels on the IQ-Raven test. So men and women are equally overconfident in this case. However, stated confidence levels are substantially higher for men when they know that there will be a tournament and that their confidence level will be observed. Men choose to enter a tournament much more frequently than women do, but this in fact corresponds to a lower stated confidence level. So women are not necessarily less competitive than men, just less confident (or possibly less strategic) in a competitive environment. It is also conceivable that women have more of an aversion to lying and misrepresentation, although there is currently no such evidence.

It is not obvious why men appear to unconsciously try to deter entrants from entering a tournament against them and women do not do so. In fact, males inflate their confidence in a tournament condition whether or not they are would-be entrants, suggesting some sort of automatic (and possibly even optimal) response to competition. An additional question is whether the behavior we observe in the voluntary-entry treatment would persist over time, as would be predicted by the equilibrium we have described.

## References


Babcock, L., G. Loewenstein, S. Issacharoff and C. Camerer (1995), "Biased Judgments of Fairness in Bargaining," *American Economic Review* **85(5),** 1337-1343.

Baumeister, R. F. (1998), "The Self," in D. Gilbert, S. Fiske and G. Lindzey (Eds.), *The Handbook of Social Psychology*, Boston: McGraw-Hill.

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

Bénabou, R. and J. Tirole (2004) "Willpower and personal rules," *Journal of Political Economy*, **112**, 848-886.

Benoit, Jean-Pierre and Juan Dubra (2011), "Apparent Overconfidence," *Econometrica*, forthcoming.

Berglas, S. and E. Jones (1978), Drug choice as a self-handicapping strategy in response to noncontingent success, *Journal of Personality and Social Psychology* **36**, 405–417.

Brunnermeier, M. and J. Parker (2005), "Optimal Expectations," *American Economic Review* **95(4),** 1092-1118.

Burks, S., J. Carpenter, L. Goette, and A. Rustichini (2009), "Cognitive abilities explain economic preferences, strategic behavior and job performance, " *Proceedings of the National Academy of Sciences* **106**, 19, 7745-7750.

Burks, S., J. Carpenter, L. Goette, and A. Rustichini (2010), "Overconfidence is a Social Signal not a Judgment Bias," mimeo.

Dana, J., R. Weber, and J. Kuang (2007), "Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness," *Economic Theory* **33**, 67-80.

DellaVigna, S. (2009), "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature* **47(2)**, 315–72.

Della Vigna, S., and U. Malmendier (2006), "Paying Not to Go to the Gym," *American Economic Review* **96**, 694-719.

Dohmen, T., and A. Falk (2006), "Performance Pay and Multi-Dimensional Sorting: Productivity, Preferences and Gender," mimeo.

Dunning, D., J.A. Meyerowitz, and A.D. Holzberg (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.

Eil, D. and J. Rao (2011), "The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself," *American Economic Journal: Microeconomics* **3(2)**, 114-138.

Ertac, S. (2008), Does Self-Relevance Affect Information Processing? Experimental Evidence on the Response to Performance and Non-Performance Feedback, mimeo**.**

Fershtman, C., U. Gneezy, and J. List (2009), “Inequity, Selfishness, and Social Norms,” mimeo.

Ferguson, M. and J. Bargh (2004), “How social perception can automatically influence behavior,” *Trends in Cognitive Sciences* **8(1)**, 33-39.

Gneezy, U., M. Niederle, and A. Rustichini (2003), “Performance in Competitive Environments: Gender Differences,” *Quarterly Journal of Economics* **118(3)**, 1049-1074.

Gneezy, U. and A. Rustichini (2004), “Gender and Competition at a Young Age,” *American Economic Review* **94(2)**, 377-381.

Grossman, Z. (2010), “Self-Signaling versus Social-Signaling in Giving,” mimeo.

Grossman, Z. (2010), “Strategic Ignorance and the Robustness of Social Preferences,” mimeo.

Grossman, Z. and D. Owens (2010), “An Unlucky Feeling: Overconfidence and Noisy Feedback,” mimeo.

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

James, W. (1884), ``What is an Emotion?'', Mind, 9, 188-205

Kartik, N. (2009), “Strategic Communication with Lying Costs,” *Review of Economic Studies* **76**, 1359-1395.

Kolditz, T. A., and R. M. Arkin (1982), “An impression management interpretation of the self-handicapping strategy,” *Journal of Personality and Social Psychology* **43**, 492–502.

Koszegi, B. (2006), “Ego Utility, Overconfidence, and Task Choice,” *Journal of the European Economic Association* **4(4)**, 673-707.

Konow, J. (2000), “Fair shares: Accountability and cognitive dissonance in allocation decisions Source, ” *American Economic Review* **90(4)**, 1072-1091.

Leary, M. (1999), “The social and psychological importance of self-esteem,” in R. M.Kowalski and M. R. Leary, eds., *The social psychology of emotional and behavioral problems: interfaces of social and clinical psychology*, 197-222.

Leary, M. R., E. S. Tambor, S. J. Terdal, and D. L. Downs (1995), “Self-esteem as an interpersonal monitor,” *Journal of Personality and Social Psychology* **68**, 518-530.

Liberman, V., S. Samuels, and L. Ross (2004), “The name of the game: predictive power of reputations versus situational labels in determining prisoner’s dilemma game moves,” *Personality and social psychology bulletin* **30**, 1175-85.

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

Milgrom, P., and J. Roberts (1982), “Limit Pricing and Entry under Incomplete Information: An Equilibrium Analysis,” *Econometrica* **50,** 443 - 459.

Möbius, M., M. Niederle, P. Niehaus and T. Rosenblat (2005-2010), “Managing Self-Confidence: Theory and Experimental Evidence, mimeo.

Moore, D. and D. Cain (2007), “Overconfidence & Underconfidence: When and Why People Underestimate (and Overestimate) the Competition,” *Organizational Behavior and Human Decision Processes* **103(2)**, 197-213.

Moore, D. and P. J. Healy (2008), “The Trouble with Overconfidence,” *Psychological Review* **115**(2), 502–517.

Niederle, M., and L. Vesterlund (2007), “Do Women Shy Away from Competition? Do Men Compete Too Much?,” *Quarterly Journal of Economics* **122(3)**, 1067-1101.

Odean, T. (1999), “Do Investors Trade Too Much?,” *American Economic Review* **89(5)**, 1279-1298.

Rabin, M. and J. Schrag (1999), “First Impressions Matter: A Model of Confirmatory Bias,” *Quarterly Journal of Economics* **114**, 37-82.

Raven, J.C. (2000), *Ravens Advanced Progressive Matrices* (APM), Pearson.

Rullière, J.L., L. Santos-Pinto, and I. Vialle (2011), Self-Confidence and Teamwork: An Experimental Test, Mimeo.

Shapira-Ettinger, K. and R. Shapira (2008), “The Constructive Value of Overconfidence,” *Review of Law & Economics* **4(3**), 751-778.

Svenson, O. (1981) Are we all less risky and more skillful than our fellow drivers?, Acta *Psychologica* **47**, 143–148.

Wolosin, R. J., S. Sherman, and A. Till (1973), Effects of Cooperation and Competition on Responsibility Attribution After Success and Failure, *Journal of Experimental Social Psychology* **9**, 220-23

# Appendix B: Instructions

The comments in square brackets are meant to illustrate instructions to the reader and were not part of the instructions.

**General instructions**

**Introduction**
Welcome to our experiment. You will receive €7 for showing up, regardless of the results. The instructions are simple. If you follow them carefully, you can earn a substantial amount of money in addition to your show up fee. Throughout the stages we will ask you to answer questions. At each stage, you will receive more detailed instructions.

You will be part of a group of 4 persons. You don't know who the other persons are, and you will remain anonymous to them. All your choices and the amount you will earn will remain confidential and anonymous, except if explicitly indicated otherwise. You will receive your earnings in an envelope. The person that puts the money in the envelopes can only see the login number that has randomly been assigned to you, and cannot match any names, student numbers, or faces with the login numbers and the decisions made.

**Payments**
There are several items in the experiment for which you can earn points. At the end of the experiment, one item is randomly chosen and your points for that item are paid in addition to the show-up fee (1 point is worth €1).

One of the participants is randomly chosen to be an assistant during the experiment. There is a random component in the experiment. The task of the assisting person will be to throw a dice which will determine the outcome.

**No deception**
Remember, we have a strict no deception policy in this lab.

**Questions**
Please remain seated and raise your hand if you have any questions, and wait for the experimenter. Please remain silent throughout the experiment.

**Part 1.**

In the first stage, all group members receive the same 15 questions. You will see a matrix with one missing segment at the bottom right. Your task is to identify the segment that would logically fit at the position of the missing segment, by choosing from the suggested answers. You can make your choice by clicking the corresponding number on the right of your screen. [A screenshot with an example question was provided.]

You can go back and forth between the questions. There is a time limit of 8 minutes. The time remaining is indicated on your screen.

After the time limit, we will rank all 4 people in your group depending on the number of questions answered correctly. The person with the highest score will get rank 1, and the person with the lowest score will get rank 4. In case of ties, the computer will randomly determine who gets the higher rank. After this, you will get some questions regarding how well you think you did.

We then randomly divide the group in 2 players A and 2 players B. Every player A will be matched against a player B. If your rank is higher than the player with which you are matched, you can receive 10 points.

**Part 2.**

All four group members have now finished with the questions, and we have determined the rank of every person.

We now ask you to indicate how likely you think that you are among the top 2 of your group. You can indicate this on a scale from 0 to 100%. Indicating 0% means that you are sure you are not among the best 2 of your group, while indicating 100% means that you are sure you are among the top 2 of your group. Similarly, 50% indicates that you think it is equally likely that you are among the best 2 of your group, or that you are not among the best 2 of your group.

We will pay you for the accuracy of your estimate. You earn more points for this item if your estimate is more accurate. The formula that is used to calculate the amount of money you earn is chosen in such a way that your expected earnings are highest when you report to us what you really believe. Reporting any value that differs from what you believe decreases your expected score for this item. If you are interested, you can find some detailed examples of this to see how this works.

[An explanation with examples was available to participants, see next page.]

**The role of player A and player B**
We matched you with one other randomly chosen person from your group. You are either Player A or B, and this is randomly determined.

[private] None of the players can see the other player's estimate of being in the top 2.

[social] Player A will not see the estimate by player B that he or she is among the best two in the group, but player B will see the estimate by Player A that he or she is among the best two of the group.

[private and social] Later on in the experiment, we will compare the rank of player A with the rank of player B, and for that item the player with the highest rank receives 10 points, the other nothing. Both of you will see who has the highest rank, and this ends the stage.

[strategic] Player A will not see the estimate by player B that he or she is among the best two in the group, but player B will see the estimate by Player A that he or she is among the best two of the group.

Later on in the experiment, after player B has observed the estimate of player A, player B will choose between two options: IN or OUT.

If player B chooses OUT, then for that item player B receives 3.5 [5.5] points and player A automatically receives 10 points. Both players will see who has the highest rank, and this ends the stage.

If player B chooses IN, we will compare the rank of player A with the rank of player B, and for that item the player with the highest rank receives 10 points, the other nothing. Both of you will see who has the highest rank, and this ends the stage.

[Participants could see their role on the next screen.]

**Determination of your score**

What follows is a brief explanation about the determination of your score, showing that it is in your interest to report truthfully what you believe in order to maximize your expected earnings.

The score is determined as follows. You start with 10 points. We subtract points depending on how close your reported belief is to the outcome. The outcome is set to 1 if you are in the top 2, and to 0 if you are not.

For instance, if you report 70% (.7), and you are in the top 2 (outcome is 1), you are .3 away from the outcome, while if you are not in the top 2 (outcome is 0), you are .7 away from the outcome.

The difference with the outcome is squared and multiplied by 10, and then subtracted from the 10 points that you start with. Thus in the example with 70%: if you are in the top 2, this gives you $10-10(.3)^2 = 9.1$. If you are not in the top 2, this gives $10-10(.7)^2 = 5.1$. You would weight these two scores by your belief about the likelihood of each occurring.

Larger differences between your reports and the outcome decrease your score proportionally more than small differences. To minimize the expected difference, and maximize your expected score, you should report what you believe.

The following examples illustrate that your expected score is highest when you report your true beliefs. All numbers used are for illustrations only and are no indication for the decisions for you to take.

**Example 1**
*You believe 50% and report 50%.* As a simple example: if you believe there is a 50% chance you are in the top 2, and you report 50%, then there is always a difference of .5 with the outcome, and since this is squared we always subtract 10 times $(0.5)^2$ points from your score, i.e. 2.5 points. Your expected score is 7.5.

*You believe 50% but you report 100%.* If you report 100%, then in one case there is no difference (if you are in the top 2) and no points are subtracted. But in the other case the difference is 1 (if you are not in the top 2), and then we subtract 10 times $10(1)^2 = 10$ from your score. If you believe the likelihood of being in the top 2 is 50%, you expect this to happen in 50% of the cases, so the amount subtracted would be 10(.5) = 5. This gives you an expected score of 5, which is lower than if you report your belief of 50%.

**Example 2**

*You believe 70% and report 70%.* As another example, suppose that you think there is a 70% likelihood that you are among the best 2. If you report 70%, your score will be either 9.1 (if you are in the top 2) or 5.1 (if you are not in the top 2). You believe that with 70% chance your score will be 9.1, and with 30% your score will be 5.1. So your expected score is $.7(9.1)+.3(5.1)=7.9$.

*You believe 70% and report 100%.* Now suppose that, instead of reporting this belief of 70%, you report another number. For instance, you report 100% (1). This means that if you are in the top 2, the outcome is as predicted, and you get $10-10(0)^2=10$ points. If you're not in the top 2, you are 1 away from the outcome, and your score will be $10-10(1)^2=0$. Since you actually expect to be in the top 2 with 70% chance, your expected score is 7. This is lower than if you would have reported 70%.

*You believe 70% and report 20%.* The same is true if you report a number below your belief, for instance 20% (.2). If you are in the top 2, your score would be $10-10(.8)^2=3.6$ points. If you're not in the top 2, your score will be $10-10(.2)^2=9.6$. Since you actually expect to be in the top 2 with 70% chance, your expected score is $0.7(3.6)+0.3(9.6)^2=5.4$, again lower than if you would have reported 70%.

The table below shows the expected scores for some more possible beliefs you may have and reports you give. As you can see, expected scores are highest when the reported belief is equal to the true belief (the cells on the diagonal that highlighted in green).

**Expected scores**

| Your report (%) | Your Belief (%) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 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 |

[private] **Part 3.**

Based on your true ranking in the group, we will send you a *report*. The report will say if you are among the two best of your group, or if you are not among the two best of your group.

However, **sometimes the report will be incorrect**. The way this works is as follows.

If you **are not** among the top two of your group, then the report will always be correct and inform you that you are not among the best two of your group.

If you **are** among the top two of your group, the report is mistaken in half of the cases. That is, in half of the cases, the report correctly informs you that you are among the top two of your group. In the other half of the cases, the report is incorrect and says you were not among the top two of your group, even if you were.

Whether or not the report you receive is correct when you are among the top two of your group, depends on the outcome of a dice throw by the assistant. You will not see the outcome, but if the assistant throws 1, 2, or 3, you will receive a correct report when you are among the top two. If the assistant throws 4, 5, or 6, you will receive an incorrect report when you are among the best two of your group. (For some groups, the incorrect report is sent after different values of the dice, but in any case the report is incorrect in half of the cases when you are among the best 2.)

After you see the report, we will ask you if you think the report is more likely to be correct or incorrect.

You earn 10 points if you are right.

[private] **Part 4**

In this part, we ask you some questions about the scenario below. The first part is always the same, but some additional information is given in the question, so please read it carefully. For this part, we randomly choose a question and this is treated as a single item.

***Scenario***

Consider two machines placed in two sides of a large production hall, left side = L and right side = R. The two machines produce rings, good ones and bad ones. Each ring that comes from the left machine, L, has a 50% chance of being a good ring and a 50% chance of being a bad ring. Each ring that comes from the right machine, R, is good. Both machines produce 100 rings every day.

The mechanic visits the production hall every day, and randomly examines one of the machines by taking one ring. On some days, he takes a ring from the left machine, and the other days he takes a ring from the right machine. Suppose the ring he takes is *good*.

We will ask you if it is more likely that the mechanic went to the left or right machine.

***Example question***

'On 50% of the days, the mechanic takes a ring from the left machine, and the other 50% of the days from the right machine.

Of the rings that come from the left machine, on average half are good and half are bad. Each ring that comes from the right machine is good.

Imagine the ring he takes is good. Is it more likely to come from the left or right machine?'

You will get 3 questions like this one. We vary the percentage of days that the mechanic goes to the left or right machine, but everything else remains the same.

You earn 10 points if you are right.

**Part 5.**

[private and social] In this part, you are informed if player A or B has the highest rank.

[strategic] Player A will not see the estimate by player B that he or she is among the best two in the group.

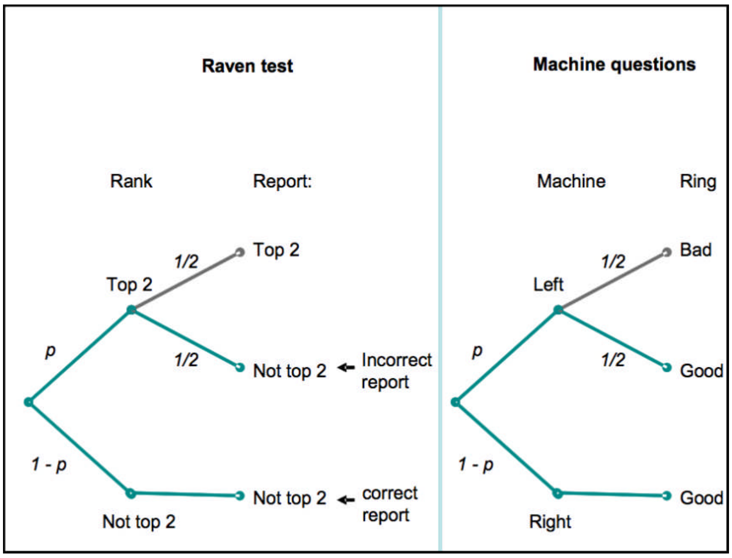
Player B will see the estimate by Player A that he or she is among the best two of the group, and then gets the choice between two options: IN or OUT.

If player B chooses OUT, then player B receives 3.5 [5.5] points and player A automatically receives 10 points. Both players will see who has the highest rank, and this ends the stage.

If player B chooses IN, we will compare the rank of player A with the rank of player B, and the player with the highest rank receives 10 points, the other nothing. Both of you will see who has the highest rank, and this ends the stage.

# Appendix C

**Figure C1: Illustration of the signal structure in the Report and the Machine question.**




This figure was not provided to subjects, and is only meant as illustration.