nature human behaviour

LETTERS

https://doi.org/10.1038/s41562-019-0666-7

Deception and self-deception

Peter Schwardmann[1*] and Joël van der Weele[2,3]

There is ample evidence that the average person thinks he or she is more skilful, more beautiful and kinder than others[1,2] and that such overconfidence may result in substantial personal and social costs[3–8]. To explain the prevalence of overconfidence, social scientists usually point to its affective benefits, such as those stemming from a good self-image or reduced anxiety about an uncertain future[9–13]. An alternative theory, first advanced by evolutionary biologist Robert Trivers[14–16], posits that people self-deceive into higher confidence to more effectively persuade or deceive others. Here we conduct two experiments (combined $n = 688$) to test this strategic self-deception hypothesis. After performing a cognitively challenging task, half of our subjects are informed that they can earn money if, during a short face-to-face interaction, they convince others of their superior performance. We find that the privately elicited beliefs of the group that was informed of the profitable deception opportunity exhibit significantly more overconfidence than the beliefs of the control group. To test whether higher confidence ultimately pays off, we experimentally manipulate the confidence of the subjects by means of a noisy feedback signal. We find that this exogenous shift in confidence makes subjects more persuasive in subsequent face-to-face interactions. Overconfidence emerges from these results as the product of an adaptive cognitive technology with important social benefits, rather than some deficiency or bias.

This study contributes to the literature on self-deception, motivated cognition and cognitive dissonance, including a growing number of studies that investigate these phenomena in economic contexts[17–21]. Our contribution is twofold. First, we show that a desire to persuade others generates overconfidence. Some previous studies provide evidence that is suggestive of this mechanism. In particular, overconfidence about performance on a cognitive test correlates with measures of social dominance[22,23]. Strategic motives such as the desire to impress an audience affect public statements of confidence[24–26], although it is not clear to what extent the communicators themselves believe their own statements. The desire to persuade others about an external event can result in biased information gathering and distorted beliefs[27]. In the study most related to ours, individuals who are primed with a desire to achieve status rate themselves higher relative to others on a series of skills and abilities relevant to attaining status in a business context[22]. Our experiment goes beyond this by studying the effect of anticipating an actual social interaction rather than the effect of a prime. Moreover, because we give participants incentives for the accuracy of their privately reported probabilistic beliefs, our results are strong evidence for self-deception.

Our second contribution is to provide rigorous causal evidence that privately held confidence about one's own performance leads to social success through both verbal and non-verbal channels. Previous work has mostly focused on the effects of stated or expressed confidence. It shows that experimental subjects have a preference for more confident probability statements by experts or advisors[28] and rate them as more convincing[29–31]. Various self-enhancement techniques are associated with higher performance in job interviews and other social situations[16,32,33]. The focus on stated confidence leaves unanswered the role of private, truly held confidence in persuasion. This is a key distinction in the evolutionary theory of Trivers, who maintains that self-deception arises from the fact that truly held confidence cannot be convincingly faked to outside observers.

Another set of papers has focused on correlations between overconfidence and social outcomes. Data from a student sample shows that overconfidence about grades is positively correlated with the estimate of one's ability by observers[34], and survey measures of self-esteem are positively correlated with job market outcomes[35–37]. But causal inferences remain unresolved in these datasets, because one cannot rule out reverse causality and unobserved covariates of confidence and social success, such as beauty or extraversion[38]. We use informative but noisy feedback to manipulate private confidence and look at its effect on persuasion in a stylized labour market interaction. In the interaction, our use of incentives for both contestants (to convince others) and evaluators (to detect true performance) simulates environments outside of the laboratory where people have money at stake[39]. This complements previous studies that manipulate confidence to investigate its effect on outcomes such as cheating[40] or status allocation by team members[22].

Our main experiment consists of two stages. In the self-deception stage, we uncover the effect of an experimentally generated social motive for self-deception on the confidence of the subject. In the deception stage, we measure the effect of higher confidence on persuasion, using exogenous variation in confidence generated by a feedback signal.

At the beginning of the self-deception stage, our 288 participants perform an intelligence test at their computer stations and receive a small remuneration depending on their performance relative to three other randomly selected subjects. We then separate subjects into a control group and a contestant group. The latter are truthfully informed that they can earn €15 later in the experiment, if they are able to persuade 'employers' of their superior performance on the intelligence test in brief face-to-face interactions. The control group does not have the opportunity to persuade employers, and receives no information about the interaction stage at this point.

Next, we elicit the private confidence of each subject about his or her performance in the intelligence test. We call this belief the 'prior belief' of a subject. Prior beliefs are elicited using an incentivized procedure based on the Becker–DeGroot–Marschak paradigm that is often used in economics[41]. Participants indicate the probability P for which they are indifferent between winning a monetary prize with probability P and winning the same prize if they are in the 'top 2', that is if their performance is among the two best performances of their randomly and anonymously selected peer group

[1]Department of Economics, University of Munich (LMU), Munich, Germany. [2]CREED, Economics Department, University of Amsterdam, Amsterdam, the Netherlands. [3]Tinbergen Institute, Amsterdam, the Netherlands. *e-mail: peter.schwardmann@econ.lmu.de

of four. After participants indicate the point of indifference in a list of probabilities, one probability is randomly drawn from the list and participants are awarded their preferred lottery for that probability. This procedure has several advantages over other elicitation mechanisms, the most important being that reporting the true subjective probability of being in the top 2 is a dominant strategy for subjects regardless of their risk preferences[42].

On the basis of Trivers' theory, we hypothesize that contestants ($n=144$), who are informed of the profitable deception opportunity, will become more confident about their performance than the uninformed control group ($n=144$). We find that the average prior beliefs among contestants (mean = 62.0%, s.d. = 17.74%) are 4.2 percentage points higher than beliefs among subjects in the control group (mean = 57.8%, s.d. = 18.4). The difference in confidence is statistically significant (two-sided t-test, $t_{286}=1.97$, $P=0.049$, $d=0.23$, 95% confidence interval (CI) = 0.012–8.39). The median and quartiles as well as the histograms of the two belief distributions are shown in Fig. 1a,b, respectively.

Table 1 presents the results of four ordinary least squares (OLS) regression models that serve to check the robustness of our treatment effect. The first column shows the raw treatment effect on prior beliefs as a percentage. In the second column, we control for individual test scores to rule out that the results are driven by differences in test scores between control and treatment group. In the third column, we add a dummy for each score as a more elaborate control for test scores. In all specifications, the statistically significant treatment effect on beliefs is roughly constant at about 4 percentage points. The positive coefficient of test scores in the second regression model tells us that beliefs correlate with actual performance and thus seem to be at least loosely tethered to reality. In the regression model in column 4, the dependent variable 'Belief bias' is a measure of individual over- and underconfidence, which is constructed by subtracting from the prior belief of a subject the average likelihood that a subject with the same test score is in the top 2 of his or her group of four. The treatment has a statistically significant effect on such overconfidence.

In Fig. 1a, a single outlier data point can be observed, which is from one subject in the control group who has a prior of 5%. When we remove this observation, the treatment effect of the main experiment is no longer significant at the 5% level (see Supplementary Results for the analysis). In light of this finding, it is reassuring that our results replicate in the follow-up experiment, as discussed below. Moreover, we note that it is not clear whether the outlier ought to be excluded in the first place. The subject in question was not in the top 2 and, therefore, rather realistic. Then, it is perfectly in line with Trivers' hypothesis that observing such realism becomes less likely when subjects are aware of a profitable opportunity to persuade.

To replicate our results and to better understand the underlying motives behind our treatment effect, we conducted a follow-up experiment with 400 subjects. We replicate our main result on self-deception. On average, participants who are informed about a profitable deception opportunity are more confident (mean = 60.8, s.d. = 18.2) than the control group (mean = 54.4, s.d. = 20.5). This difference is statistically significant (two-sided t-test, $t_{398}=2.75$, $P=0.0063$, $d=0.32$, 95% CI = 1.79–10.87). Further analyses are provided in the Supplementary Results. We also perform a mini meta-analysis by pooling the two samples. The treatment comparison is statistically significant (two-sided t-test, $t_{686}=3.9033$, $P<0.001$, $d=0.31$, 95% CI = 2.97–8.98). The corresponding equivalent of regression in Table 1 yields a statistically significant treatment dummy in all specifications (see Supplementary Table 3).

At the end of the follow-up experiment, we conducted several additional elicitations to further investigate the motivations behind self-deception. Notably, we aim to understand whether the desire to be more persuasive is a (subconscious) driver for self-deception. To this end, we incentivize subjects to state their belief that 'confidence pays'—that is, that a subject with higher confidence in the main experiment was more likely to get a higher evaluation. We find suggestive evidence that a persuasion motive is indeed at work. We also explicitly test for several alternative motives for self-deception, such as a wish to reduce one's guilt about lying, anticipatory utility concerns and a desire to appear consistent in front of the experimenter, and find no evidence for these. The Supplementary Results provide a detailed analysis of the follow-up experiment.

Table 1 | The effect of strategic motives on confidence

Dependent variable	Prior belief	Prior belief	Prior belief	Belief bias
Contestant (d)	4.201	3.770	4.098	5.648
P	0.049	0.045	0.027	0.019
95% CI	0.0117–8.391	0.0824–7.458	0.462–7.734	0.945–10.35
Score		1.149		−3.289
P		0.000		0.000
95% CI		0.908–1.391		−3.671 to −2.908
Dummy for each score	No	No	Yes	No
Constant	57.81	42.32		52.04
P	0.000	0.000		0.000
95% CI	54.80–60.83	37.99–46.65		45.28–58.81
Observations	288	288	288	288
R^2	0.013	0.241	0.938	0.602

OLS regressions of the determinants of private prior belief (column 1–3), and our measure of overconfidence (column 4). The dependent variable, Prior belief, is measured as the privately stated probability of being in the top 2 of one's group of four (scale, 0–100). Belief bias is a measure of overconfidence (see text). Contestant (d) is a dummy variable to indicate the treatment. Score is the performance on the test.

In the deception stage of the main experiment, we tested for the strategic benefits of higher confidence through its effect on persuasiveness. The 144 participants who were in the control group during the self-deception stage now take the role of employer and, in face-to-face interactions, interview the 144 participants in the role of contestant. We chose to make the interactions face-to-face rather than using written messages because Trivers' theory assigns an important role to physical ticks or give-away tells associated with low confidence. During the interactions, contestants are incentivized to convince employers that they did well on the intelligence test, whereas employers have an incentive to detect those contestants that truly were the best performers.

The interactions followed a speed-dating protocol. Equipped with a pen and evaluation sheets, employers leave their computer stations and sit down in front of the contestants. A group of four employers is matched with a group of four contestants. There are four rounds of interviews so that each of the four employers in a group interviews each of the four contestants in a group. Each interview takes place behind a partition to assure some level of privacy. On the ring of a bell, contestants say one sentence: 'I believe that my performance was in the top 2 of my group with ... per cent probability'. In the blank, each contestant verbally fills in a number between 0 and 100.

After the interviews, employers return to their computer stations and enter their evaluations as well as the verbal messages of the contestants into the computer. Each employer states the probability

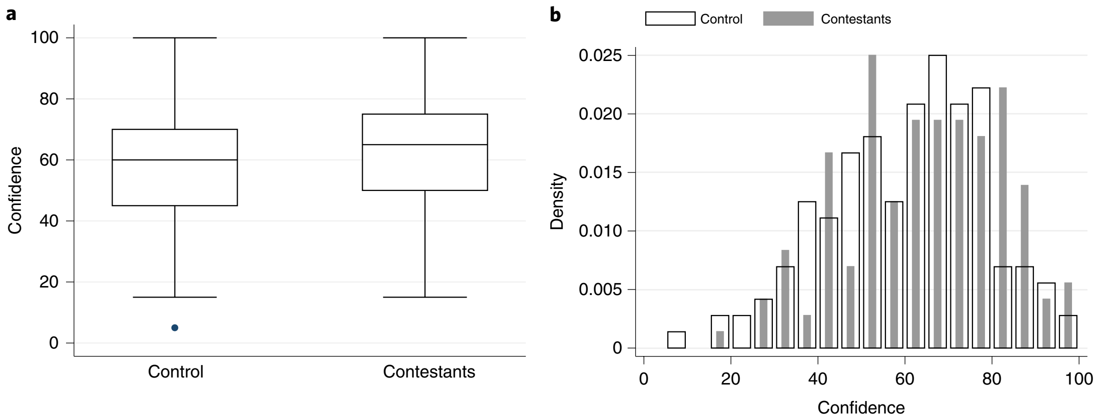

Fig. 1 | Prior beliefs of participants in treatment and control groups. a, Tukey box plots of prior beliefs of participants in the treatment group (contestants, $n = 144$) and the control group ($n = 144$). The box shows the second and third quartiles, separated by the median; whiskers cover all data within the first and fourth quartiles (within 1.5× the interquartile range); the dot indicates the single outlier observation. The *y* axis is the subjective probability of being in the top 2 of the group (scale, 0–100). **b**, Histogram of prior beliefs of participants in the treatment group (contestants, $n = 144$) and the control group ($n = 144$). The *x* axis is the subjective probability of being in the top 2 of the group (scale, 0–100).

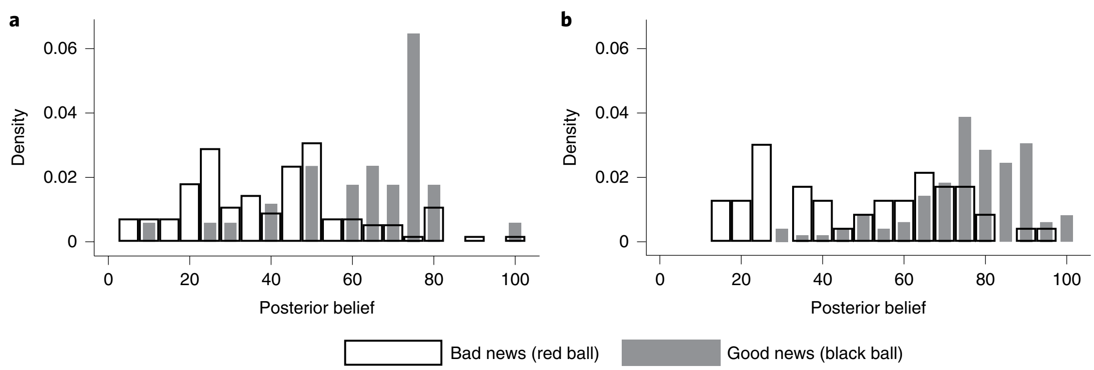

Fig. 2 | Histograms of posterior beliefs for those not in the top 2 and those in the top 2. a, Posterior beliefs of participants who were not in the top 2 ($n = 144$). **b**, Posterior beliefs of participants in the top 2 ($n = 144$). The *x* axis is the subjective probability of being in the top 2 of the group (scale, 0–100).

that they associated with each of the four contestants that they interviewed being in the top 2 of the contestant group. One of these evaluations is selected at random and the employer is paid according to the Becker–DeGroot–Marschak mechanism described above, with a chance to win €10 in case of a favourable outcome. A contestant is paid €15 if he or she received one of the two best evaluations by a randomly selected employer and nothing otherwise.

Before the interviews, every contestant receives noisy feedback about his or her relative performance in the test and then again privately states his or her belief about how likely it is that he or she was in the top 2 of his or her group of four. We refer to the belief we elicit at this time as the 'posterior belief' of a subject. Regarding the noisy feedback, contestants are truthfully told that the computer draws a ball from one of two virtual urns that contain 20 balls of two different colours. If their performance was actually in the top 2, then the ball would come from an urn with 15 black balls and 5 red balls. If their performance was not in the top 2, then the ball would come from an urn with 15 red balls and 5 black balls. Thus, although a black ball is good news about performance, the signal is noisy and some subjects randomly receive a ball of the wrong colour. It is this variation that we use to identify the effect of (over)confidence on persuasiveness. The merits of this design choice compared to other possible designs are discussed further in the Supplementary Discussion.

Figure 2 shows the histograms of post-feedback beliefs for each performance level (top 2 and not top 2), separated by whether a subject received good or bad news. The signal has a strong effect on the beliefs of the contestants, which makes sense, given that it is informative. Within each panel, the difference between the grey and white bars shows how the noise component or error in the signal causes subjects with a similar performance to adopt different beliefs. This purely random shift in confidence is uncorrelated with any personal characteristics and can be exploited to estimate the causal effect of confidence on evaluations. However, such causal inference also requires that the feedback affects evaluations only through its effect on confidence. In order to not violate this condition, the minimalist nature of the interaction between employers and contestants assures that contestants are not able to communicate their signal.

We hypothesize that increased confidence makes participants more persuasive in their attempts to convince others of their superior performance. In the Supplementary Results, we consider the relationship between private confidence and communication strategies. Here, our main focus lies on the role of confidence in determining the evaluations of employers. To test our hypothesis, we use a series of linear regression models. The first four regression models in Table 2 focus on our baseline treatment. The OLS regression in column 1 features only a correlation. It shows that more confident contestants—that is, those with a higher posterior belief—receive

Table 2 | Effect of confidence on employer evaluations

	OLS regression		IV regression		
Dependent variable	Evaluation	Evaluation	Evaluation	Evaluation	Evaluation
Posterior	0.221		0.320	0.290	−0.252
P	0.002		0.012	0.028	0.160
95% CI	0.0898–0.352		0.0713–0.570	0.0322–0.548	−0.603–0.0994
Round	−0.792	−0.792	−0.792	−0.917	−2.676
P	0.541	0.541	0.525	0.463	0.090
95% CI	−3.430–1.846	−3.433–1.850	−3.232–1.649	−3.365–1.531	−5.771–0.418
Black ball (d)		5.650			
P		0.022			
95% CI		0.876–10.42			
Top 2		3.907	−1.295	−1.864	13.78
P		0.349	0.788	0.687	0.002
95% CI		−4.554–12.37	−10.73–8.141	−10.94–7.209	4.877–22.68
Message				0.284	0.518
P				0.002	0.001
95% CI				0.103–0.465	0.224–0.813
Constant	39.02	47.38	33.85	15.13	26.93
P	0.000	0.000	0.000	0.003	0.000
95% CI	28.73–49.32	39.06–55.70	21.05–46.66	5.219–25.04	14.20–39.66
Condition	Baseline	Baseline	Baseline	Baseline	LD
Observations	384	384	384	384	192
R^2	0.070	0.043	0.057	0.112	0.243

Determinants of employers' evaluations of contestants. The first two columns show OLS regressions; the last three columns show instrumental variable (IV) regressions, in which the posterior belief is instrumented by 'Black ball'. The dependent variable, Evaluation, is the employer's stated belief that the contestant is in the top 2 (scale, 0–100). Posterior is the privately stated confidence of the contestant of being in the top 2 after receiving the feedback signal (scale, 0–100). Black ball is a dummy that indicates whether the contestant received positive feedback. Top 2 is a dummy that indicates whether the contestant was in the top 2 performers of his or her group of four. Message is the belief stated by the contestant to employer in the interview (scale, 0–100). Round indicates the interview round (scale, 1–4). Standard errors are clustered at the subject's group of four contestants and four employers. LD, lie detection.

higher evaluations from employers. The regression in column 2 of Table 2 tells us that contestants who receive a good signal obtain higher evaluations, controlling for whether the contestant actually was in the top 2. This demonstrates that confidence has a positive causal effect on evaluations by employers.

To obtain a more concrete estimate of the size of the causal effect of confidence on evaluations, we carry out an instrumental variable estimation known as two-stage least squares. In the first stage, we quantify the variation in posterior beliefs that is induced by our 'instrument', the noise component of the feedback signal. This first-stage regression is reported in the Supplementary Table 5 and shows that the instrument explains a substantial part of posterior beliefs. We then take the variation in beliefs explained by the instrument in the first stage, and use it to estimate the effect of beliefs on evaluations in the second stage. To make sure that these estimates are based only on random variation in beliefs, both first and second stage regressions include the actual performance, that is, whether the subject was in the top 2, as a control variable.

The results of the second stage regressions are reported in column 3 of Table 2. A 1 percentage point increase in the posterior belief leads to a 0.32 percentage point increase in the evaluation that a contestant receives. The fact that the coefficient in the instrumental variable regression exceeds the coefficient in the OLS regression is indicative of the OLS regression being biased by endogeneity. We cannot be sure about the exact reason; one possible cause is an omitted personality trait, such as arrogance, that is positively correlated with confidence and negatively correlated with evaluations. Although we do not measure the arrogance of our subjects, we do measure the trait 'assertiveness' on the 'Big Five Aspect Scales' in the extroversion domain[43]. As described in Supplementary Table 6, we find that there is a negative relationship between assertiveness and evaluations, but the small correlation is not significant and cannot explain the difference between the OLS and instrumental variable results.

To investigate how much of the effect of beliefs on evaluations works through the verbal messages that contestants send to employers, we add messages to the set of regressors in column 4. Messages are a statistically significant predictor of evaluations and when we include them into the instrumental variable regression, the coefficient of posterior beliefs decreases slightly but remains statistically significant. These results suggest that higher confidence leads to higher evaluations through both verbal and non-verbal channels. Quantitatively, the two channels appear to be of roughly equal importance. In Supplementary Table 6, we establish that these results are robust to more flexible model specifications for messages and the inclusion of control variables related to characteristics of the contestant, the employer and other contestants in the same group.

The model in column 5 of Table 2 presents the deteminants of the evaluations of the employers in a lie-detection treatment. In contrast to our baseline condition, employers in the lie-detection treatment followed a small lie-detection tutorial, based on psychological research on indicators that have diagnostic power for lies[44,45]. We hypothesize that the effect of the confidence of the contestants is larger in a situation in which employers have a higher ability to spot lies. The literature on lie detection generally finds that subjects in the laboratory are not great at spotting lies in others, although they do slightly better than chance[16,39]. However, these studies may understate lie-detection abilities outside of the laboratory, where we seek to deceive spouses, professional recruiters and loan officers. The lie-detection condition was designed to erode some of the artificial advantage bestowed on the deceiver by the laboratory context. Unfortunately, owing to a randomization failure in the feedback signal in the lie-detection treatment, very few contestants received a signal that did not align with their actual performance. This makes it hard to distinguish between actual performance and beliefs. Moreover, we find that our instrument for beliefs in the lie-detection condition is rather weak (see the Supplementary Table 5, column 3). As a result, it is difficult to interpret these data, and we merely report them here for the sake of full disclosure. Taken at face value, the results in column 5 of Table 2 would suggest that employers in the lie-detection treatment are not responsive to the confidence of the contestants at all, but are able to identify true ability. This would be surprising, because spotting true ability is different from spotting lies, defined as increases of stated confidence over truly held confidence, which is what employers in the condition were trained to detect. The Supplementary Results provide further analyses of the lie-detection treatment and establish robustness of our results to pooling data from the lie-detection and baseline treatments.

Our findings are succinctly summarized in the words of Mark Twain, who writes in his autobiography: “When a person cannot deceive himself the chances are against his being able to deceive other people”. We find that confidence management is an adaptive cognitive technology, enabling trade-offs between the benefits of self-enhancement and realism. Overconfidence represents an optimal response to environments in which social influence and persuasion are often crucial[14,46,47]. One implication of our findings is that overconfidence is likely to be more prevalent in settings in which its strategic value is highest, that is, in cases in which measures of true ability are noisy, competition is fierce and persuasion is an important part of success. It may arise in employer–employee relationships because of its strategic benefits in job interviews and wage negotiations. Arguably, confidence may be even more valuable among the self-employed, whose economic survival often depends more immediately on persuading investors and customers. We would also expect overconfidence to be rife amongst high-level professionals in finance, law and politics.

Methods

Ethical approval was obtained from the University of Amsterdam research priority area Behavioral Economics. The experiments took place at the Munich Experimental Laboratory for Economic and Social Sciences (MELESSA). Subjects were randomly recruited from the MELESSA subject pool using the recruitment system ORSEE[48]. The experiments were programmed in *z*-tree[49] and random assignment to groups and conditions within the experiments was implemented using the random-number-generating function in *z*-tree. Both experiments lasted slightly over 1 h and the average subject earned €16.45 (minimum, €4; maximum, €28.5).

Our main experiment featured 288 participants, divided over 18 sessions of 16 subjects each. Overall, 61% of the participants was female and the average age was 23.1 years. We ran a first set of sessions, featuring 192 subjects, in March 2015. The number of subjects was determined on the basis of budgetary constraints and a single pilot with 32 subjects. In October 2016, we ran an additional 6 sessions, comprising 96 subjects, as a robustness check requested by a referee. All of the results in the combined sample were already obtained in data from the first set of sessions. The sample size for the follow-up experiment was determined by recruiting the maximum number of subjects that remained in the MELESSA subject pool, as estimated by the laboratory manager. This resulted in a total sample of 400 subjects. Of the total pool, 59% of the participants was female and the average age was 24.6 years. No participants were excluded from the analysis for either experiment.

All statistical tests in our regressions use threshold levels of $\alpha = 0.05$, $\alpha = 0.01$ and $\alpha = 0.001$, and we report P values for all of our main results. For the t-tests of our deception stage results, we test for normality of the samples using the Shapiro–Francia test. We cannot reject normality at the 5% for any of our data, unless otherwise reported. All reported tests are two-sided. No data points were excluded from the analysis. Data collection and analysis were not performed blind to the conditions of the experiments.

Detailed procedures. After coming into the laboratory, participants spent 15 min on an introduction to the Becker–DeGroot–Marschak mechanism we that used to elicit beliefs. The mechanism requires participants to indicate which probability P makes them indifferent between winning a monetary prize with probability P and winning the same prize when an uncertain event E occurs. After participants indicated the point of indifference in a list of probabilities, one probability is randomly drawn from the list and participants are awarded their preferred lottery for that probability. We told subjects that reporting the true subjective probability of E maximizes their chance of winning the prize. In contrast to other elicitation mechanisms, the incentives for truthful reporting do not depend on risk preferences. The veracity of this claim has been established previously by different authors[42]. We only proceeded once all subjects correctly answered four control questions about the belief elicitation.

The main experiment had two parts, which we refer to as the ‘self-deception stage’ and the ‘deception stage’. In the self-deception stage, we investigated whether the announcement of a profitable deception opportunity had an influence on the beliefs of the participants about their own performance in an intelligence test. To this end, we separated the group into contestants, who were informed about the later deception opportunity, and a control group of participants, who were not. In the deception stage, we investigated whether (over)confidence made the contestants more persuasive. Contestants competed to persuade employers of their superior performance on the intelligence test in a face-to-face interaction. Employers were the participants who constituted the control group in the self-deception stage of the experiment. The sequencing of experimental tasks is depicted in Supplementary Table 1. Translated experimental instructions can be found online at https://osf.io/v2r3q/. The design of the self-deception stage of the follow-up experiment is identical. The deception stage of the follow-up experiment was slightly different and is described in the Supplementary Information.

Self-deception stage. Participants were then divided into anonymous groups of four and proceeded to the intelligence test. The test consisted of 15 Raven matrices of varying difficulty and participants had 10 min to solve as many as they could. Participants were awarded two points for each correct answer and lost one point for each incorrect or omitted answer. The subjects with the two top scores in their anonymous group of four earned €2. Their earnings, and hence their ranking within the group, were only communicated to them at the end of the experiment.

We administered our main treatment after the intelligence test. Of the four groups of four in each session, two groups were assigned to the role of contestants and two groups were designated as controls. While the control group was not told anything about the deception stage at this point, the contestants received a short summary of the instructions for the deception stage, which are reproduced in the Supplementary Methods.

Next, we used the Becker–DeGroot–Marschak mechanism to elicit the prior belief of each subject about the probability that he or she is among the top 2 performers in his or her group of four. We refer to this event as ‘top 2’. The prize for the elicitation was €3. The instructions made it very salient that elicited beliefs were strictly confidential and would never be shown to another subject.

After participants submitted their prior beliefs, we gave them noisy feedback on their performance. Participants were told that they would be shown a ball drawn from one of two virtual urns, which contained 20 balls of different colours. If their performance was actually in the top 2, the ball would come from an urn with 15 black balls and 5 red balls. If their performance was not in the top 2, the ball would come from an urn with 15 red balls and 5 black balls. Therefore, a black ball constituted good news about their performance. After subjects had observed the ball, they reported their posterior belief about being in the top 2.

Deception stage. For the face-to-face interactions, the 16 subjects in a session were divided into two groups, each of which consisted of four contestants and four employers. The latter formed the control group in the first stage of the experiment. Before the interviews began, employers were given an evaluation sheet for each contestant on which to write down the message of the contestant and their evaluations of the contestant’s relative performance, honesty, confidence, likability and attractiveness. The interviews followed a speed-dating protocol. Employers left their computer stations and sat down in front of the contestants. There were four rounds of interviews so that every employer would get to interview each of the four contestants in the same group.

On the ring of a bell, contestants said one sentence: ‘I believe that my performance was in the top 2 of my group with ... per cent probability’. In the blank, each contestant verbally filled in a number between 0 and 100 that we refer to as a contestant’s message. During the interviews, none of the 144 contestants said anything more than this sentence. After the sentence was said, there were a few seconds in which employers could scrutinize contestants’ faces and body language, before the bell rang again to mark the end of a round. Employers were given time to fill in their evaluation sheets, before moving on to the next contestant.

After the four rounds of interviews, employers returned to their computer stations and entered their evaluations as well as contestants’ messages into the computer. Each employer had to state the probability he or she associated with each of the contestants being in the top 2. The computer program enforced that the four elicited probabilities, elicited as percentages between 0 and 100, added up to 200, so employers had to revise their original evaluations if necessary. One of these evaluations was selected at random and the employer was paid for it according to the Becker–DeGroot–Marschak mechanism described above, with a chance to win €10 in case of a favourable outcome. The contestant was paid €15 if he or she was among the two best evaluated contestants by the employer in the selected round, and nothing otherwise.

At the same time, contestants participated in a task that elicited their degree of lying aversion[50]. During this task, participants are asked to imagine themselves in the position of a company CEO, who can earn money by deceiving shareholders about the value of the company. Again using a Becker–DeGroot–Marschak mechanism, participants are asked to indicate repeatedly whether they would report deceptively or not, for which the amount of money they could earn from doing so was increased in four steps from 0 cents to 120 cents. After seeing their pay-offs in the experiment, all subjects filled out a questionnaire about their background characteristics and their assertiveness.

In one third of the sessions, employers participated in a lie-detection tutorial before they embarked on the interview. The tutorial featured on-screen written instructions and lasted 3 min. The tutorial used four lie-detection tips from the wiki http://www.wikihow.com/Detect-Lies on how to recognize 'tells' associated with lying, namely fidgeting, face-touching, fast breathing and incongruent facial expressions. In addition, it explained that avoidance of eye contact is an unreliable indicator of lying. Contestants in the sessions of the lie-detection treatment did not know about the tutorial.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

We report that all of the experimental data and all of the experimental conditions performed in this research project are included in the paper and its Supplementary Information. All data and codes to reproduce the analysis are available from the Open Science Framework (https://osf.io/k6hy5).

Code availability

Our STATA Do-file for the data analysis is downloadable from the Open Science Framework at https://osf.io/k6hy5.

Received: 14 November 2018; Accepted: 21 June 2019;
Published online: 29 July 2019

References

1. Alicke, M. D. & Sedikides, C. Self-enhancement and self-protection: what they are and what they do. *Eur. Rev. Soc. Psychol.* **20**, 1–48 (2009).
2. Baumeister, R. F. in *The Handbook of Social Psychology* (eds Gilbert, D. et al.) (McGraw-Hill, 1998).
3. Barber, B. M. & Odean, T. Boys will be boys: gender, overconfidence, and common stock investment. *Q. J. Econ.* **116**, 261–292 (2001).
4. Glaser, M. & Weber, M. Overconfidence and trading volume. *Geneva Risk Insur. Rev.* **32**, 1–36 (2007).
5. Biais, B., Hilton, D., Mazurier, K. & Pouget, S. Judgemental overconfidence, self-monitoring, and trading performance in an experimental financial market. *Rev. Econ. Stud.* **72**, 287–312 (2005).
6. Malmendier, U. & Tate, G. Who makes acquisitions? CEO overconfidence and the market's reaction. *J. Financ. Econ.* **89**, 20–43 (2008).
7. Koellinger, P., Minniti, M. & Schade, C. 'I think I can, I think I can': overconfidence and entrepreneurial behavior. *J. Econ. Psychol.* **28**, 502–527 (2007).
8. Dawson, C., de Meza, D., Henley, A. & Arabsheibani, G. R. *The Power of (Non) Positive Thinking: Self-employed Pessimists earn more than Optimists* IZA Discussion Paper No. 9242 (Forschungsinstitut zur Zukunft der Arbeit Institute for the Study, 2015).
9. Kunda, Z. The case for motivated reasoning. *Psychol. Bull.* **108**, 480–498 (1990).
10. Taylor, S. E. & Brown, J. D. Illusion and well-being: a social psychological perspective on mental health. *Psychol. Bull.* **103**, 193–210 (1988).
11. Brunnermeier, M. K. & Parker, J. A. Optimal expectations. *Am. Econ. Rev.* **95**, 1092–1118 (2005).
12. Köszegi, B. Ego utility, overconfidence, and task choice. *J. Eur. Econ. Assoc.* **4**, 673–707 (2006).
13. Bénabou, R. Groupthink: collective delusions in organizations and markets. *Rev. Econ. Stud.* **80**, 429–462 (2013).
14. Trivers, R. in *Social Evolution* 395–420 (Benjamin/Cummings, 1985).
15. Trivers, R. *The Folly of Fools: The Logic of Deceit and Self-Deception in Human Life* (Basic Books, 2011).
16. Von Hippel, W. & Trivers, R. The evolution and psychology of self-deception. *Behav. Brain Sci.* **34**, 1–16 (2011).
17. Babcock, L., Loewenstein, G., Issacharoff, S. & Camerer, C. F. Biased judgments of fairness in bargaining. *Am. Econ. Rev.* **85**, 1337–1343 (1995).
18. Konow, J. Fair shares: accountability and cognitive dissonance in allocation decisions. *Am. Econ. Rev.* **90**, 1072–1091 (2000).
19. Di Tella, R., Pérez-Truglia, R., Babino, A. & Sigman, M. Conveniently upset: avoiding altruism by distorting beliefs about others. *Am. Econ. Rev.* **105**, 3416–3442 (2015).
20. Bénabou, R. & Tirole, J. Mindful economics: the production, consumption, and value of beliefs. *J. Econ. Perspect.* **30**, 141–164 (2016).
21. Gino, F., Norton, M. I. & Weber, R. A. Motivated Bayesians: feeling moral while acting egoistically. *J. Econ. Perspect.* **30**, 189–212 (2016).
22. Anderson, C., Brion, S., Moore, D. A. & Kennedy, J. A. A status-enhancement account of overconfidence. *J. Pers. Soc. Psychol.* **103**, 718–735 (2012).
23. Burks, S. V., Carpenter, J. P., Goette, L. & Rustichini, A. Overconfidence and social signalling. *Rev. Econ. Stud.* **80**, 949–983 (2013).
24. Ewers, M. & Zimmermann, F. Image and misreporting. *J. Eur. Econ. Assoc.* **13**, 363–380 (2015).
25. Charness, G., Rustichini, A. & van de Ven, J. Self-confidence and strategic behavior. *Exp. Econ.* **21**, 72–98 (2018).
26. Thoma, C. Under- versus overconfidence: an experiment on how others perceive a biased self-assessment. *Exp. Econ.* **19**, 218–239 (2016).
27. Smith, M. K., Trivers, R. & von Hippel, W. Self-deception facilitates interpersonal persuasion. *J. Econ. Psychol.* **63**, 93–101 (2017).
28. Keren, G. & Teigen, K. H. Why is $p = .90$ better than $p = .70$? Preference for definitive predictions by lay consumers of probability judgments. *Psychon. Bull. Rev.* **8**, 191–202 (2001).
29. Brewer, N. & Burke, A. Effects of testimonial inconsistencies and eyewitness confidence on mock-juror judgments. *Law Hum. Behav.* **26**, 353–364 (2002).
30. Radzevick, J. R. & Moore, D. A. Competing to be certain (but wrong): market dynamics and excessive confidence in judgment. *Manage. Sci.* **57**, 93–106 (2011).
31. Tenney, E. R., Meikle, N. L., Hunsaker, D., Moore, D. A. & Anderson, C. Is overconfidence a social liability? The effect of verbal versus nonverbal expressions of confidence. *J. Pers. Soc. Psychol.* **116**, 396–415 (2019).
32. Barrick, M. R., Shaffer, J. A. & DeGrassi, S. W. What you see may not be what you get: relationships among self-presentation tactics and ratings of interview and job performance. *J. Appl. Psychol.* **94**, 1394–1411 (2009).
33. Sedikides, C., Hoorens, V. & Dufner, M. in *Self-concept, Motivation and Identity: Underpinning Success with Research and Practice* (eds Guay, F. et al.) 29–55 (Information Age Publishing, 2015).
34. Lamba, S. & Nityananda, V. Self-deceived individuals are better at deceiving others. *PLoS One* **9**, e104562 (2014).
35. Drago, F. Self-esteem and earnings. *J. Econ. Psychol.* **32**, 480–488 (2011).
36. de Araujo, P. & Lagos, S. Self-esteem, education, and wages revisited. *J. Econ. Psychol.* **34**, 120–132 (2013).
37. Kaniel, R., Massey, C. & Robinson, D. T. *The Importance of Being an Optimist: Evidence from Labor Markets.* National Bureau of Economic Research Working Paper Series No. 16328 (National Bureau of Economic Research, 2010).
38. Möbius, M. M. & Rosenblat, T. S. Why beauty matters. *Am. Econ. Rev.* **96**, 222–235 (2006).
39. Belot, M. & Van de Ven, J. How private is private information? The ability to spot deception in an economic game. *Exp. Econ.* **20**, 19–43 (2017).
40. Aronson, E. & Mettee, D. R. Dishonest behavior as a function of differential levels of induced self-esteem. *J. Pers. Soc. Psychol.* **9**, 121–127 (1968).
41. Becker, G. M., DeGroot, M. H. & Marschak, J. Measuring utility by a single-response sequential method. *Behav. Sci.* **9**, 226–232 (1964).
42. Schlag, K. H., Tremewan, J. & an der Weele, J. J.. A penny for your thoughts: a survey of methods for eliciting beliefs. Exp. Econ. 18, 457–490 (2015).
43. DeYoung, C. G., Quilty, L. C. & Peterson, J. B. Between facets and domains: 10 aspects of the Big Five. *J. Pers. Soc. Psychol.* **93**, 880–896 (2007).
44. Ekman, P. & Friesen, W. V. Nonverbal leakage and clues to deception. *Psychiatry* **32**, 88–106 (1969).
45. Zuckerman, M., DePaulo, B. M. & Rosenthal, R. in *Advances in Experimental Social Psychology* Vol. 14 (ed. Berkowitz, L.) 1–59 (Academic, 1981).
46. McKay, R. T. & Dennett, D. C. The evolution of misbelief. *Behav. Brain Sci.* **32**, 493–510 (2009).
47. Mercier, H. & Sperber, D. Why do humans reason? Arguments for an argumentative theory. *Behav. Brain Sci.* **34**, 57–74 (2011).
48. Greiner, B. An online recruitment system for economic experiments. *Forsch. Wiss. Rechn.* **63**, 79–93 (2003).
49. Fischbacher, U. z-Tree: Zurich toolbox for ready-made economic experiments. *Exp. Econ.* **10**, 171–178 (2007).
50. Gibson, R., Tanner, C. & Wagner, A. F. Preferences for truthfulness: heterogeneity among and within individuals. *Am. Econ. Rev.* **103**, 532–548 (2013).

Acknowledgements

P.S. and J.v.d.W. acknowledge support from the German Science Foundation (DFG) through grant CRC TR190. J.v.d.W. acknowledges support from the Dutch National Science Foundation (NWO) through a personal VIDI grant (452-17-004). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. We thank K. Barron, T. Buser, S. Chassang, A. Coutts,

B. Enke, B. Kőszegi, Y. Le Yaouanq, G. Loewenstein, J. Maier, D. de Meza, T. Murooka, K. Schmidt, P. Seabright, J. van de Ven, A. Wuppermann, F. Zimmermann and several seminar audiences for helpful comments. J. Hannane provided research assistance.

Author contributions

P.S. and J.v.d.W. conducted the experiments and contributed to the conception, design, analysis and writing of the paper in equal parts.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41562-019-0666-7.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to P.S.

Peer review information: Primary Handling Editor: Aisha Bradshaw

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Peter Schwardmann

Last updated by author(s): Jun 18, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Data were collected on computer terminals, using the software z-Tree, which is available free of charge to academic researchers from Prof. Urs Fischbacher at Konstanz University.
Data analysis	Data analysis was performed with the software package STATA (version 15).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

We report all of the experimental data and all of the experimental conditions performed in this research project. All data and codes to reproduce the analysis are available at Open Science Framework https://osf.io/k6hy5/.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This is a behavioral experiment, yielding quantitative data.
Research sample	In total 688 people participated. Participants are undergraduate students at the University of Munich's MELESSA experimental laboratory. The average reported age was 24 (s.d. 6.2, minimum 16, maximum 65). In total, 410 women and 278 men participated.
Sampling strategy	In March 2015, we ran a first set of sessions for the main experiment featuring 192 subjects. The number of subjects was determined on the basis of budgetary constraints and a single pilot with 32 subjects. In October 2016, we ran an additional six sessions, comprising 96 subjects as requested by a referee. All of the results in the combined sample were already obtained in data from the first set of sessions. For the follow-up experiment, we chose to use the maximum number of remaining subjects still available in the subject pool. According to the lab manager, we could recruit 400 participants, which is our sample size.
Data collection	Data were collected on the computer, using the software z-Tree (see above).
Timing	We piloted our design with 32 subjects in December 2014. We then ran 18 sessions of 16 subjects each in March 2015. In October 2016, on the request of referees for an earlier submission, we ran a further 6 sessions to replace data from six sessions that were not to be featured in the main analysis (see data exclusions). Our second experiment was run in October 2016.
Data exclusions	We report all experimental data and all experimental conditions in the paper. No data points were excluded from the paper. One experimental condition does not feature in the main analysis. In this additional condition we warned contestants about the lie-detection tutorial of the employers. The condition pertains to a separate research question about the effects of warnings. We do not analyze this condition in this paper, because several commentators as well as previous referees argued it was outside the scope of this paper, and advised/requested us to take it out. For robustness and to maintain power, we ran an additional 6 sessions of 16 subjects each (see above) to replace the data from the warning condition in the main analysis of the first experiment.
Non-participation	No participants dropped out or declined participation.
Randomization	Subjects were randomly assigned to conditions.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☒	☐	Antibodies
☒	☐	Eukaryotic cell lines
☒	☐	Palaeontology
☒	☐	Animals and other organisms
☐	☒	Human research participants
☒	☐	Clinical data

Methods

n/a	Involved in the study	
☒	☐	ChIP-seq
☒	☐	Flow cytometry
☒	☐	MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	See above
Recruitment	Subjects were recruited from the MELESSA subject database using the ORSEE software developed by Ben Greiner, which randomly selects subjects from the database to receive an invitation.

Ethics oversight	University of Amsterdam, Research Priority Area Behavioral Economics.

Note that full information on the approval of the study protocol must also be provided in the manuscript.