

Joy Leads to Overconfidence, and a Simple Remedy

Philipp D. Koellinger and Theresa Treffers

Author Note

Philipp D. Koellinger, Amsterdam Business School, University of Amsterdam, Plantage Muidergracht 12, 1018 TV Amsterdam, The Netherlands, T: +31 20 525 4171, P.D.Koellinger@uva.nl *and* Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands.

Theresa Treffers, School of Industrial Engineering & Innovation Sciences, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, T: +31 40 247 5532, T.Treffers@tue.nl (Corresponding author).

Abstract

Overconfidence has been identified as a source of suboptimal decision making in many real-life domains, and it often has far-reaching consequences. Here, we demonstrate a causal mechanism that leads to overconfidence and show a simple, effective remedy for it in an incentive-compatible experimental study. We show that joy induces overconfidence if the reason for joy (an unexpected gift) is unrelated to the judgment task and if participants were not made specifically aware of our mood manipulation. In contrast, we observed well-calibrated judgments among participants in a control group who were in their resting mood. Furthermore, we found well-calibrated judgments among participants who received the joyful mood induction together with questions that forced them to reflect on their current mood, its cause, and the (ir)relevance of its cause to our judgment tasks. Our findings suggest that being aware of one's positive mood and the reason for that mood can be an effective short-term remedy for overconfidence.

Keywords: Joy; overconfidence; mood; emotion; judgment

Joy Leads to Overconfidence, and a Simple Remedy

Overconfidence is “widespread, stubborn, and costly” (Kahneman, 2011, p. 257). For example, entrepreneurs are often overconfident about their chances of success and their own skills compared to others, which often results in suboptimal returns on their investments (e.g., Cooper, Woo, & Dunkelberg, 1998; Camerer & Lovallo, 1999; Koellinger, Minniti, & Schade, 2007; Simonsohn, 2010). Financial investors, both amateur and professional, “stubbornly believe that they can do better than the market, contrary to an economic theory that most of them accept and contrary to what they could learn from a dispassionate evaluation of their personal experience” (Kahneman, 2011, p. 217; Barber & Odean, 2001; Kirchler & Maciejovsky, 2002). Bettors who gamble on football games prefer low-probability wagers – and frequently lose – because they have too much confidence in their own expertise (Golec & Tamarkin, 1995). Inventors also seem to be more overconfident and optimistic compared with the general population. They often continue to invest substantial amounts of time and money into their ideas despite strongly supported advice to cease their efforts – not surprisingly, the end results are mostly disappointing (Åstebro, 2003; Åstebro, Jeffrey, & Adomza, 2007). Managers, planners, and even academics often ignore statistical information about how long it will take them to complete large projects. Instead, they rely on their own expertise and insights, which often results in unrealistic plans, unexpected delays, higher costs and disappointing outcomes (Merrow, Phillips & Myers, 1981; Kahneman & Lovallo, 1993).

Overconfident CEOs tend to overestimate the returns on their investment projects and fail to reduce their personal exposure to company-specific risk (Malmendier & Tate, 2005). They frequently overpay for target companies and undertake value-destroying mergers (Malmendier & Tate, 2008). Nevertheless, overconfidence may help managers become promoted to CEO in the first place (Goel & Thakor, 2008).

Examples of overconfidence are not just limited to the economic and social science spheres: Medical studies have shown that physicians' overconfidence is likely to be an important reason for severe diagnostic errors that result in increased patient morbidity and mortality (Berner & Graber, 2008). In 2008, the *American Journal of Medicine* devoted an entire special issue to overconfidence and clinical decision making (See Issue 5) and cited overconfidence as one of the most common reasons for medical errors among physicians.

In some cases, overconfidence is even a matter of war and peace: Rational states should not go to war against one another without thinking they can win - and clearly, at least one side must be wrong (Johnson, 2009). The ability to avoid bad judgments and decisions that are induced by overconfidence is therefore important (e.g., Dunning, Heath, & Suls, 2004; Ely, Graber, & Croskerry, 2011; Kuhn, 2002; Norman & Eva, 2010).

Overconfidence

Confidence judgments can be miscalibrated in three different ways (Moore & Healy, 2008). First, individuals may be inaccurate in the precision of their judgment (i.e., overprecision, e.g., Kirchler & Maciejovsky, 2002). Second, people may be imprecise in the judgment of their own skills and abilities (i.e., *absolute* over- or underconfidence) (Blavatsky, 2009; Taylor & Brown, 1988). Third, people may be biased in their judgment of themselves relative to others (i.e., *relative* over- or underconfidence) (Larrick, Burson, & Soll, 2007; Moore & Small, 2007). In the present study, we investigate the influence of joy on *absolute* and *relative* overconfidence in one's own performance on a general-knowledge task. The distinction between judgments of absolute and relative performance is important because self-reported judgments are psychometrically different from comparative judgments (Vautier & Bonnefon, 2008), and the mental processes that lead to these judgments may differ (Urbig, Stauf, & Weitzel, 2009).

Previous studies have suggested that overconfidence is influenced by gender differences (e.g., Barber & Odean, 2001; Bengtsson, Persson, & Willenhag, 2005) and differences in personality (Schaefer, Williams, Goodie, & Campbell, 2004). Furthermore, twin studies show a moderate heritability of overconfidence (Cesarini, Johanneson, Lichtenstein, & Wallace, 2009), suggesting that some people are more prone to overconfidence for (yet unknown) biological reasons. However, gender, personality, and heritability estimates do not explain the underlying mechanisms that lead to overconfidence, nor do they suggest potential remedies for it.

Earlier studies have shown that overconfidence is a context- and task-dependent phenomenon. For example, absolute overconfidence appears to be greatest during difficult tasks, whereas relative overconfidence appears to be greatest during easy tasks (Hoelzl & Rustichini, 2005; Larrick et al., 2007; Moore & Healy, 2008; Moore & Small, 2007). These inconsistencies may be explained by the fact that knowledge about one's own performance is uncertain, but knowledge about the performance of others is even more uncertain. Consequently, people's post-task self-assessments are regressive, and their estimates of others are even more regressive (Moore & Healy, 2008). Again, however, these insights do not help people avoid overconfidence during a given decision-making situation.

Potential Remedies against Overconfidence

Dunning et al. (2004) review several methods that may be used in specific situations to avoid overconfidence or at least to limit its damage-causing potential. For example, engineers calculating the amount of concrete needed to hold a dam or a bridge may benefit from multiplying their calculations by a safety factor of 3 or even 8. Similarly, managers can add "buffer time" into their budgets when they plan projects.

Furthermore, some empirical evidence suggests that skeptical feedback from knowledgeable persons who are not personally involved in a project can help to prevent

overconfidence to some extent (Dunbar, 1995; Heath et al., 1998). More generally, confronting people with objective risk information may reduce the bias in their judgments (Weinstein, 2003). Kahneman and Lovallo (1993) argued that overconfidence may be prevented by focusing on an “outside view” that looks beyond the details of the specific situation and instead examines statistical information about a group of cases that are similar in relevant respects. For example, an aspiring restaurant owner should not ask how likely it is that her particular business will succeed. Instead, she should ask what the average survival rate and financial performance of all restaurants in the relevant area have been in the past few years. This outside perspective could provide some protection against forecasts that are not within the realm of reasonable possibilities. However, objective information about statistical risk is often difficult to obtain, and in many cases, it is arguable what the relevant reference category for a specific situation should be. This limitation is particularly relevant for situations that involve a high degree of novelty (Dunning et al., 2004).

Another possible remedy against overconfidence has been suggested by Gary Klein (Kahneman 2011, pp. 255-265), who proposed a planned “premortem” to avoid the overconfidence trap. The idea is that a group of people responsible for making a particular decision should meet just before making the final commitment to a project. At the meeting, they should stage a hypothetical event one year in the future in which they discuss why the project has miserably failed. The possible effectiveness of the “premortem” rests on unleashing the imagination of knowledgeable individuals and on legitimizing doubts about the project within the group of decision makers. Although we are not aware of any empirical evidence that the “premortem” works, it may be effective for group decision making.

The decision-making situation we study in the present article resembles that faced by individuals who need to make accurate judgments about their own performance without feedback or any other outside information that they could use to anchor their beliefs.

Accordingly, the remedies against overconfidence discussed above are not available to our decisions makers. Instead, their only source of information is their own knowledge and experiences and the experimental environment. In such situations, people's judgment may be influenced by their moods and emotions. Indeed, an extensive literature has identified the crucial influence of affective states on judgment and on decision making more generally (e.g., Clore & Huntsinger, 2007; Forgas, 1995; Isen & Geva, 1987; Schwarz & Clore, 1983), and some scholars have also started investigating these relationships in the brain (e.g., Bechara, Damasio, & Damasio, 2000). We refer to this literature and link it to overconfidence as a prominent judgment bias. A similar approach was taken in a recent study by Ifcher and Zarghamee (2011), who found that mild positive affect increases overconfidence. We attempt to go a step further in our study by aiming to identify a potential mechanism for this result that suggests a possible remedy.

Affect and judgment

Judgments have been shown to be directly (Schwarz & Clore, 1983) or indirectly (Bower, 1981) influenced by affect¹ in an affect-congruent manner. Typically, people's resting mood is slightly positive (Schwarz & Clore, 1983), and the cause of this base-level positive mood is typically diffuse and not salient to the individual (Clore & Huntsinger, 2007). Studies that do not make the cause of a given mood salient typically report affect-congruent judgments in various contexts, such as life satisfaction (Schwarz & Clore, 1983), consumer products (Yeung & Wyer, 2004), culpability (Kadous, 2001), risk estimates

¹ "Affect" is often used as an umbrella term for moods and emotions. Whereas moods can be defined as diffuse, objectless affective states, emotions are regarded as affective states that are typically directed toward a certain object and that manifest in particular action tendencies, expressive reactions, and physiological changes. Like emotions, moods are evaluative states characterized by positive and negative feelings. In this study, we use the term "mood" instead of "emotion" because we believe that the definition of mood fits our participants' baseline affect and our gift-based induction procedure better than the definition of emotion does.

(Lerner, Gonzalez, Small, & Fischhoff, 2003), and estimates of physical space (Proffitt, 2006).

Affect-congruent influences on judgments can be explained by the concept-priming hypothesis (Bower 1981; Isen, Shalcker, Clark, & Karp, 1978) and the affect-as-information hypothesis (Schwarz & Clore, 1983). Although these hypotheses present different, but possibly parallel, mechanisms for how affect impacts judgments, both state that current affective states result in affect-congruent outcomes.

The concept-priming hypothesis suggests that affect may prime (i.e., activate in memory) material that is affect-congruent, leading to affect-congruent judgments (Bower 1981). Happy moods may facilitate the recall of happy memories and inhibit the recall of sad memories, and vice versa (Isen et al. 1978). These affect-congruent memories could lead to affect-congruent judgments and decision outcomes.

In contrast, the affect-as-information hypothesis proposes that affective cues of moods and emotions directly influence judgments by providing experiential and bodily information that is attributed to the object of judgment (Clore & Huntsinger, 2007; Schwarz & Clore, 1983). Thus, individuals may ask themselves, “How do I feel about it?” In doing so, they may mistake pre-existing feelings for a reaction to the target. Therefore, positively valenced affect could encourage individuals to continue, thereby providing them with a “go” signal that may lead to affect-congruent judgments, such as an increase in confidence.

Building on the affect-as-information hypothesis, studies have shown that the misattribution of irrelevant causes of affective states to judgment objects that leads to affect-congruent judgments may weaken or disappear entirely if the irrelevant source of the affect is salient (e.g., Gasper, 2004; Gasper & Clore, 2000; Kadous, 2001; Schwarz & Clore, 1983). For example, Schwarz and Clore (1983) found that the impact of negative moods on life satisfaction and happiness was eliminated when the irrelevant source of the affective

information (i.e., the weather) was made explicit. When participants were provided with an explanation for their current affective state that was irrelevant to their evaluation of their lives, they were less likely to use their affective cues as an informational basis for their evaluative judgments.

The Present Research

Our study examines whether the attribution of affective states to irrelevant causes, as affect-as-information theory suggests, is applicable to judgments about one's own performance and may result in overconfidence. Based on the assumption of affect-congruent judgments, we test whether joy can be a transient cause of overconfidence. Furthermore, our goal is to demonstrate that awareness of the irrelevance of the cause of joy to the judgment task serves as a practical, easy tool against overconfidence.

To measure overconfidence, we employed a general knowledge task of medium difficulty that can be expected to have a relatively high number of well-calibrated judgments compared with very easy or very difficult tasks (Hoelzl & Rustichini, 2005; Larrick et al., 2007; Moore & Healy, 2008; Moore & Small, 2007). Although this decreases our a priori chance of observing overconfidence and puts an upper boundary on the possible effect sizes of the remedies against overconfidence that we study, we believe that most of the relevant decision-making situations we named above fall into the medium difficulty category. It also makes our study comparable that of Ifcher and Zarghamee (2011) and our own previous work (see Appendix 1). Furthermore, we use an incentive-compatible method (Smith, 1982) to elicit confidence judgments that further improve internal and external validity because the precision of judgments has been shown to depend on whether the decision-maker faces real (financial) consequences (Blavatsky, 2009; Cesarini, Sandewall, & Johanneson, 2006; Hoelzl & Rustichini, 2005).

Prior considerations

To aid the design of our experiment, we attempted to gauge the plausible effect sizes of experimental mood manipulations on overconfidence by consulting prior literature and by conducting an exploratory pretest at a large public university in Germany (see Appendix 1 for details about the pretest). The study by Ifcher and Zarghamee (2011) was the only existing study we could identify that induced positive affect and measured the effect of this mood manipulation on overconfidence. We computed the effect sizes that the experimental treatment had on absolute and relative overconfidence in Ifcher and Zarghamee (2011) using the reported regression coefficients and their standard errors. The positive mood induction had direct effects that varied between *Cohen's* $d = 0.22$ and $d = 0.62$, depending on the model specification. The effects of the positive mood induction on absolute and relative overconfidence in our own pretest varied between $d = 0.28$ and $d = 0.63$, also depending on the model specification. The sample-size weighted average effect size observed in these studies was $d = 0.42$.

<Insert Table 1 here>

Our own pretest and the study by Ifcher and Zarghamee (2011) induced joy using film clips combined with questionnaires about the participants' moods that were administered either before and after the overconfidence measure (in our pretest) or only after the measure (in Ifcher and Zarghamee, 2011). The impact of the joy induction on both overconfidence measures was diametrically opposite in these two studies. This led us to hypothesize that the influence of positive affect on overconfidence critically depends on whether the irrelevant source of positive affect is salient to the participants.

We designed our current study to specifically test this assumption, which was not possible in previous experiments. We expected somewhat larger effect sizes within our main experiment ($E(d) \approx 0.5$) than in the previous studies because we aimed to isolate the effect

that we thought caused the large differences between the results of our own pretest and those of the study by Ifcher and Zarghamee (2011).

Based on our expected effect size of $d \approx 0.5$ and practical considerations, we decided at the outset to terminate data collection after one week in the laboratory with the aim of recruiting at least 50 participants for each of four experimental groups. A priori power calculations showed that with $N_{\text{group}} = 50$ and $d \approx 0.5$, we would have 80% power to find between-group differences in means with $p = 0.05$ in a one-sided test (using G*Power 3.1.7, Faul, Erdfelder, Buchner, & Lang, 2009). We conducted our main experiment in a different university and country from our pretest and the study by Ifcher and Zarghamee (2011) to ensure the independence of observations. Data analyses started after data collection was finished.

Materials and Methods

Participants

Our main experiment was conducted at a large public university in the Netherlands. It included 226 participants randomly distributed across four groups.² The sample consisted of 124 male and 102 female management and economics students with an average age of 21 years ($SE = 0.16$). The youngest and oldest participants were 18 and 29 years old, respectively. The participants were informed that they would participate in an experiment on decision making that had real financial payoffs. They completed the experiment in a cubicle without the ability to communicate with other participants, and there were no time restrictions for completing the experiment (average completion time was 35 minutes).

² Because of the large pool of 2,500 participants from which we recruited, most of the participants did not know any other participants in the experiment. We conducted various statistical tests and found no date or time effects in the participants' behavior.

Procedure and experimental treatments

Upon arrival, the participants drew a number that assigned them to a cubicle in the laboratory. The participants signed consent forms confirming that they had read and understood the terms and conditions of the experiment and that the experimenters had adequately answered all of their questions. In each session, there was one control group and three treatment groups.

Before the experiment started, half of the participants received a small bag of gummy bears as an unexpected gift (X_1). The gummy bears were always handed out by the same experimenter, who entered the cubicles with a smile and the words, “This is a thank-you gift for your participation. The experiment will start in a few minutes.” The gifts were distributed so that the other participants did not notice. Gift-related mood induction processes have been regularly and successfully used to induce positive mood states (Gerrards-Hesse, Spies, & Hesse, 1994; Isen, Daubman, & Gorgoglione, 1987; Isen & Geva, 1987; Westermann, Spies, Stahl, & Hesse, 1996). We used the gift for two different purposes, which resulted in two different treatment groups. First, we used the gift to amplify participants’ positive resting mood, without making the purpose of this intervention obvious. This treatment represents the *joy* group (see the experimental design in Table 2).

Second, we used the gift to provide participants in the *joy awareness* group with a salient cause for their joyful mood that was clearly irrelevant to the judgment tasks that would follow. To make the irrelevant source (i.e., the gift) of this mood manipulation evident, the participants in the *joy awareness* group answered a short scale (X_2) after they received the gummy bears and before the experiment began. The participants indicated their answers about four short statements on a 10-point Likert-type scale with the poles 1 “I strongly disagree” and 10 “I strongly agree” (Gaspar, 2004; Gaspar & Clore, 2000). The statements were as follows: “1. I am currently in a joyful mood” ($Mean = 7.6$); “2. The present I

received for participating in this experiment had a positive influence on my mood” ($Mean = 8.0$); “3. The fact that I received a present does not influence how much I know about science and culture” ($Mean = 8.9$); and “4. The fact that I received a present should not influence my ability to make good judgments” ($Mean = 8.5$). Responses to the first two items indicate a positive effect of the gift on participants’ mood state.

One-quarter of the participants received a treatment that made them aware of the influence of overconfidence on people’s decisions at the beginning of the experiment, without receiving a gift. Our *overconfidence awareness* group received treatment in the form of a short text (X_3) stating that overconfidence is a widespread and notorious bias that negatively influences most people’s judgments and decisions. The text further stated that participants could maximize their payoff in this experiment if they were aware of this bias and if they judged their performance accurately. We included this treatment because the control group in our pretest (see Appendix 1) was clearly overconfident, and we were interested in determining whether informing people about the overconfidence bias would also lead to better-calibrated judgments.

The *control* group in our main experiment did not receive a gift or a particular pretext, but started directly on the overconfidence measure (O_1). For all other groups, O_1 followed the experimental treatment(s) (Table 2). The instructions for the experimental task (see Overconfidence measure) stated that the participants’ payoff at the end of the experiment would depend on their performance on the subsequent task. The participants received an attendance fee of 4 EUR and the opportunity to earn up to 12 EUR more.

After the experimental task, the participants answered standard questions (O_2) about their age, gender, current occupation, and level of education. The participants also completed

short measures of personality (Gosling, Rentfrow, & Swann, 2003),³ risk preferences (Dohmen, Falk, Huffman, Schupp, Sunde, & Wagner, 2005), general self-efficacy (Schwarzer & Jerusalem, 1995; Cronbach's $\alpha = 0.76$), and optimism (Scheier, Carver, & Bridges, 1994; $\alpha = 0.74$). Table 2 presents an overview of the experimental design.⁴

At the end of the experiment, payoffs were determined by the computer according to the experimental instructions and observed behavior. The participants received their payoffs separately and sequentially after they signed individual receipts. On average, the participants earned 7.30 EUR ($SE = 0.29$); 34 participants earned nothing above their participation fee, 98 participants won an additional 6 EUR, and 81 participants received 12 EUR in addition to their participation fee.

Overconfidence Measure

To elicit the participants' beliefs about their absolute and relative performance, we used the incentive-compatible experimental design developed by Urbig et al. (2009) and a general knowledge quiz with ten multiple-choice questions of medium difficulty (Blavatskyy, 2009). In Urbig et al.'s (2009) method, the participants' payoffs depended on their absolute performance on the general knowledge quiz and their choices between two different payoff

³ Internal reliabilities for measurements of extraversion (Cronbach's $\alpha = 0.63$), agreeableness ($\alpha = 0.01$), conscientiousness ($\alpha = 0.60$), emotional stability ($\alpha = 0.69$), and openness ($\alpha = 0.41$) were not satisfactory. Hence, we used the ten personality items for a rotated factor analysis and extracted four components with eigenvalues that were greater than 1, namely, extraversion, emotional stability, conscientiousness, and openness. Agreeableness was not found to be a separate factor and was therefore not included in the analyses. The principal component analysis (PCA) considers all of the available information, whereas the sum scores of the items for one personality dimension ignore the personality differences among participants. The factor scores and sum scores were almost perfectly correlated for all four of the five examined personality dimensions ($r > 0.88$, $p < 0.001$).

⁴ The experimental design was adapted and extended from the pretest (see Appendix 1).

mechanisms. Specifically, each participant completed two sets of ten binary payoff choices. One set elicited absolute confidence, and the other set elicited relative confidence. We randomized the order of these sets to prevent ordering effects.

<Insert Table 2 here>

In both sets of binary choices, the participants decided which of two payoff mechanisms they preferred. One of the two payoff mechanisms was a random lottery with a known probability of winning. The probabilities of winning increased in ten steps from 5% to 95% from a choice one to ten for both sets. The alternative payoff mechanism depended on how many questions the participants answered correctly in the general knowledge quiz. In the measure of absolute overconfidence, this payoff mechanism generated a win if one randomly chosen quiz question was answered correctly. In the measure of relative overconfidence, the payoff mechanism generated a win if another randomly chosen participant in the experiment answered *fewer* questions correctly. In the case of a draw between both participants, a fair coin flip decided the winner.

The participants did not receive feedback about their performance on the quiz. Therefore, they had to exercise careful judgment about their performance to choose the payoff mechanism that maximized their expected payoff. The switching point from the performance-based payoff mechanism to the lotteries with known probabilities was used to compute the levels of absolute and relative overconfidence.

This method has a number of important advantages. First, rather than directly asking participants for their level of confidence, their beliefs are elicited in a strictly incentive-compatible experiment (Smith, 1982) that strongly reinforces deliberate behavior and precise judgment (see Urbig et al., 2009 for the proof). This is important because appropriate financial incentives improve the precision of a judgment (Cesarini et al., 2006; Fischhoff, Slovic, & Lichtenstein, 1977; Hoelzl & Rustichini, 2005) and influence behavior in risky

situations (Holt & Laury, 2002; 2005). Second, the measure allows absolute and relative overconfidence to be differentiated within subjects. In our experiment, absolute overconfidence refers to overestimating the number of correctly answered quiz questions, whereas relative overconfidence refers to overestimating one's own performance in answering the quiz questions compared to a randomly chosen opponent. Third, the measure differentiates between people with well-calibrated judgments and those without and allows the degrees of overconfidence to be specified. Fourth, Urbig et al.'s (2009) method avoids confounding risk attitudes with overconfidence (Golec & Tamarkin, 1995). Finally, the method provides unambiguous monetary incentives for participants to maximize the number of correctly answered quiz questions because their chance of winning strictly increases with their performance in the quiz. This is important because we are interested in examining the participants' ability to judge their own maximum performance, not their ability to optimize the predictability of their performance (Budescu, Wallsten, & Au, 1997; Moore & Healy, 2008).

Based on these data, six measurements were generated (see Urbig et al., 2009). *Absolute performance (ap)* was measured as the percentage of correctly answered quiz questions. *Relative performance (rp)* took a value of 1 if the participant performed better than a randomly assigned participant, 0 otherwise, and 0.5 in the case of a draw between the two participants. *Confidence (c)* was measured by the switching point from the payoff mechanism based on *absolute* performance to the lottery with known probabilities. For example, if a participant chooses the absolute performance payoff when the alternative is winning with 45% or less but otherwise chooses the lotteries that have a probability of winning 55% or more, this implies that the participant believes that he or she answered 50% of the quiz questions correctly ($c = 0.5$). *Relative confidence (rc)* was measured by the switching point from the payoff mechanism based on *relative* performance to the lottery with known

probabilities. For example, if a participant chooses the relative performance payoff when the alternative is winning with 65% or less but otherwise chooses the lotteries that have a 75% or higher probability of winning, this implies that the participant believes that his or her performance ranks in the top 30% of all people who participated in the experiment ($rc = 0.7$). *Absolute overconfidence* (oc) is the difference between absolute confidence and absolute performance, $oc = c - p$. *Relative overconfidence* (roc) is the difference between relative confidence and relative performance, $roc = rp - rc$. Both oc and roc can vary between -1 and 1, where a value of 0 indicates a well-calibrated judgment of absolute and relative confidence.

We calculated roc based on comparison with the median performance in the experiment rather than a randomly chosen opponent. This procedure removes the noise resulting from the random matching of participants from the relative overconfidence measure roc .⁵ The measurements of absolute and relative overconfidence can vary between -1 and 1, where a value of 0 indicates a well-calibrated judgment of absolute and relative confidence.

Results

On average, the participants answered 4.6 ($SE = 0.13$) out of ten questions correctly. The median performer answered five questions correctly. The best and the worst performances were ten ($N_{10} = 2$) and 0 ($N_0 = 1$) questions answered correctly, respectively. An analysis of variance (ANOVA) revealed no statistical differences between the experimental groups in their performance ($F[3, 209] = 0.32, p = 0.81, partial \eta^2 = 0.01$) or

⁵ The random matching procedure implies that a participant would sometimes be classified as over- or underconfident depending on the randomly assigned opponent. Because the participants could not guess who their opponent would be, the best they could do was to form an opinion about how good the median participant in this experiment would be and whether they believed themselves better than the median.

their performance range (test of homogeneity of variances, Levene's $F(3, 209) = 0.57, p = 0.63$).

Figure 1 shows the distribution of well-calibrated and under- and overconfident participants in our sample. On average, self-confidence judgments were better calibrated for absolute ($ap_{well-calibrated} = 0.18, SE = 0.03$) than for relative ($rp_{well-calibrated} = 0.06, SE = 0.03$) performance. Underconfidence occurred equally for absolute ($ap_{underconfidence} = 0.43, SE = 0.04$) and relative performance ($rp_{underconfidence} = 0.42, SE = 0.04$). Overconfidence was more common for relative performance ($rp_{overconfidence} = 0.52, SE = 0.01$) than for absolute performance ($ap_{overconfidence} = 0.39, SE = 0.02$).

<Insert Figure 1 here>

Correlation results (Appendix 3) show that absolute and relative overconfidence tended to co-occur in our data (*Pearson's* $r = 0.70, p = 0.00$) and that they were negatively associated with participants' performance on the quiz (oc: $r = -0.58, p = 0.00$; roc: $r = -0.78, p = 0.00$). Furthermore, absolute ($r = 0.16, p = 0.02$) and relative ($r = 0.16, p = 0.02$) overconfidence were more frequently observed in the *joy* treatment. In addition, absolute overconfidence was weakly correlated with the *joy awareness* treatment ($r = -0.12, p = 0.09$). The overconfidence measures and some of the treatment groups were also weakly correlated with risk preferences, personality traits, and general self-efficacy. Furthermore, the age and employment status of the participants were not equally distributed across the treatment groups. We controlled for these potential confounds in the multivariate analyses below.

A descriptive comparison of absolute overconfidence across the experimental groups (Figure 2) shows that participants in the *joy* treatment group ($M = 0.05, SE = 0.03$) made overconfident judgments, whereas participants in the *control* group ($M = -0.03, SE = 0.03$) and the *joy awareness* group ($M = -0.04, SE = 0.03$) judged their performances as slightly

underconfident. People in the *overconfidence awareness* group ($M = 0.004$, $SE = 0.03$) judged their absolute performance very accurately.

The effect of the *joy* treatment on absolute overconfidence compared with the *control* group was $t(102) = 1.87$, $p = 0.07$, $d = 0.37$. *Joy awareness* decreased absolute overconfidence compared with the *joy* group, with $t(106) = 2.33$, $p = 0.02$, $d = 0.45$. A one-sample, one-sided t -test shows that the participants in the *joy* treatment were marginally overconfident in an absolute sense ($t_{mean=0}(52) = 1.56$, $p_{one-sided} = 0.06$, $d = 0.43$).

<Insert Figure 2 here>

Descriptive comparisons for relative overconfidence (Figure 3) indicate that the participants in the *joy* treatment group ($M = 0.16$, $SE = 0.06$) again showed the highest level of overconfidence, followed by the *control* group ($M = 0.05$, $SE = 0.06$). Both awareness groups (i.e., the *joy awareness* group [$M = -0.01$, $SE = 0.06$] and the *overconfidence awareness* group [$M = 0.01$, $SE = 0.06$]) showed well-calibrated confidence judgments about their relative performance.

The effect of the *joy* treatment on relative overconfidence compared with the *control* group was $t(102) = 1.47$, $p = 0.15$, $d = 0.29$. Compared with the *joy* treatment group, *joy awareness* decreased relative overconfidence, with $t(106) = -2.06$, $p = 0.04$, $d = 0.40$. The one-sample, one-sided t -test shows that the participants in the *joy* treatment group were on average overconfident in a relative sense ($t_{mean=0}(52) = 2.87$, $p_{one-sided} = 0.005$, $d = 0.80$).

<Insert Figure 3 here>

Based on the observed effect sizes, post-hoc power calculations show that we had 74% power to find the effect of *joy awareness* compared with *joy* on absolute overconfidence and 66% power to find the effect on relative overconfidence (1-sided tests with $p = 0.05$ using G*Power 3.1.7, Faul et al., 2009).

To check the robustness of our findings, we ran an analysis of variance (ANOVA) and four analyses of covariance (ANCOVA) using absolute and relative overconfidence as dependent variables and the experimental groups as the between-subjects factor.⁶

<Insert Table 3 here>

Discussion

Our study identifies joy as a relevant transient cause of absolute and relative overconfidence. Participants who received an unexpected gift (a small bag of gummy bears; e.g., Isen et al., 1987, Isen & Geva, 1987) at the beginning of the experiment (i.e., the *joy* group) were more overconfident in their judgments about their absolute and relative performance on a general knowledge quiz than were participants in the control group who did not receive such a gift. If a small present such as gummy bears can cause an increase in overconfidence, other, more intense joyful experiences may have even stronger and more enduring effects. For example, Kurtz, Wilson and Gilbert (2007) found that a gift from an unknown source created longer lasting positive feelings than did an equivalent gift from a known source.

Our finding that joy leads to overconfidence is consistent with current theories that suggest that positive affect leads to affect-congruent judgments. While we argue that our findings may be best explained by the affect-as-information mechanism (Clore & Huntsinger, 2007; Schwarz & Clore, 1983), we cannot exclude the possibility that parts of our findings may also be explained by the activation of affect-associated memories, as suggested by the concept-priming hypothesis (Bower, 1981; Isen et al., 1978). Affect-as-information proposes that individuals use their pre-existing feelings as an informational cue for the decision at hand, leading to affect-congruent judgments. Under this assumption, biased judgments and suboptimal decisions (e.g., overconfidence) occur if the reason for the positive affect (e.g., an

⁶ All results are robust to the inclusion of the 13 initially excluded participants ($N = 226$, see also Table 2).

unexpected gift, sunny weather, a funny movie) is not salient to the decision maker and is irrelevant to the judgment task. This biasing effect on judgments, which is caused by using affect as information, can be eliminated if the irrelevant source of the current affective state is salient (e.g., Gasper, 2004; Gasper & Clore, 2000; Kadous, 2001; Schwarz & Clore, 1983).

Our study demonstrates this effect for absolute and relative overconfidence and thus suggests a simple remedy against overconfidence: if decision makers are aware of their current moods and consciously reflect on the (possibly irrelevant) reasons for their moods, they may be able to avoid misattributing their moods to their decisions and may consequently improve the quality of their judgments. If the indirect concept-priming mechanism had applied instead of affect-as-information, our remedy would not have been successful. Under concept-priming assumptions, subjects who are made aware of their affective misattribution would still use their positive affect for the retrieval of affect-related memories, leading to overconfidence. In contrast, affect-as-information is more likely to explain our results: the subjects' awareness of the fact that their good mood was caused by the unexpected gift led to better-calibrated judgments.

However, we do not know how long these improvements in judgment accuracy last. We also do not know what effects joy might have if it is induced by a relevant source. Under such circumstances, it may be possible that joy (or any other affective state) increases the quality of judgments. Future research could investigate how the relevance of an affective source influences individual's judgment and decision-making quality. Furthermore, our findings were related to a general-knowledge task context in a tightly controlled experimental setting. Testing whether our remedy against overconfidence also works in different settings would be a worthwhile area for future research. Because our participants were management and economics students, our results may be interpreted as the current behavioral pattern of future managers and entrepreneurs.

We did not expect to find that the participants in the control group were underconfident in an absolute sense and only slightly overconfident in a relative sense. In fact, our pretest (see Appendix 1) found that the participants in the control group were on average overconfident to a degree comparable to that of the *joy* group in the main experiment reported above. The PANAS-X measures used in our pretest showed that the participants in the control group were in a positive resting mood (see Appendix 1). We suspect that the slightly different experimental environments contributed to this observed difference: The pretest in Germany was conducted in a behavioral laboratory that had 24 computers that were located in a large room and were only separated by blinds between the tables. In addition, the room had windows through which incoming sunlight may have affected participants' affective states and their choices in the experiment. In contrast, the participants in the main experiment in the Netherlands were seated in small, closed booths without sunlight that did not allow the participants to see or hear any activities around them. We suspect (but did not measure) that being seated in such an isolated, closed booth had an undesirable and slightly negative influence on the subjects' mood, which may in turn have influenced their confidence. Furthermore, the main experiment was conducted during the fall in the Netherlands, when the weather is typically grey and rainy, compared with summer in southern Germany, when the weather is generally warm and sunny.

Because we did not observe overconfidence in the control group of our main experiment, we cannot tell whether the *overconfidence awareness* treatment is an effective remedy against overconfidence. We can only observe that mean judgments in the *overconfidence awareness* group were well calibrated on average. It may be promising to investigate in further studies whether awareness of a bias alone can improve the quality of judgments and decisions.

The average effect sizes in our experiment comparing the *joy* and *joy awareness* groups ($\text{Avg}(d) = 0.43$ in Table 3) were slightly below our (over)optimistic expectation of $d \approx 0.5$. Although this outcome led to a minor loss of achieved statistical power (see Results section) relative to our expectations, our study still fairs relatively well compared with most other experiments in psychology, which have an average power of only 0.5 (Maxwell, 2004).

We observed that our participants' average judgment precision was lower for relative overconfidence than for absolute overconfidence. This outcome was reflected in standard errors that were twice as large on average for relative overconfidence compared with absolute overconfidence (Figures 2 & 3) and a higher share of people who were relatively overconfident (Figure 1). A plausible explanation for this is the higher degree of uncertainty that people face when they compare themselves to others. Relative judgments of performance require both awareness about oneself and an estimate about the performance of others (Moore & Healy, 2008). However, our suggested remedy against overconfidence was equally effective in both scenarios. Specifically, our *joy awareness* group did not show evidence of absolute or relative overconfidence, while the *joy* group did.

An important question that our study does not address is whether accurate, realistic judgments are always desirable. Optimistic self-delusion may be a positive causal factor that helps people cope with the challenges of life and persevere in the face of difficulty (Busenitz & Barney, 1997; Hoffrage, 2004; Seligman, 1991; Taylor, Kemeny, Reed, Bower, & Gruenewald, 2000). Indeed, some studies have reported positive associations between CEOs' overconfidence and firms' innovative performance (Galasso & Simcoe, 2011; Hirshleifer, Low, & Theoh, 2012). Similarly, more (over)confident entrepreneurs also self-report that they are more innovative than their less-confident colleagues (Koellinger, 2008). The direction of causality is, however, unclear in all of these studies.

Despite the methodological challenges of identifying the effect of overconfidence on performance in field data, we acknowledge that overconfidence may indeed have partially positive effects: Theoretical work suggests that populations with a small share of overconfident individuals may actually have evolutionary advantages compared to populations with entirely unbiased perceptions (Bernardo & Welch, 2001).

Indeed, Taylor and Brown (1988, 1994) have argued in two widely cited theoretical studies that optimistic self-delusion may even promote mental health, happiness, the ability to care for others, and the capacity for creative and productive work. Others have disagreed with this hypothesis and found evidence that overconfidence is related to problems with social interactions (Colvin & Block, 1994; Colvin, Block, & Funder, 1995) and depression (Dunning & Story, 1991; Fu, Koutstaal, Fu, Poon, & Cleare, 2005), contradicting the view that overconfidence is a positive aspect of mental health. As far as empirical studies find a positive relationship between joy, happiness, or subjective well-being and overconfidence (Freedman, 1978; Cummins & Nistico, 2002), our study suggests that the direction of causality may be opposite to what Taylor and Brown (1988, 1994) advocate for transient states: Rather than overconfidence promoting happiness, overconfidence may be the result of positive affect that the decision-maker has mistakenly misattributed as relevant information.

To the extent that overconfidence can be regarded as negative (such as in the examples we listed at the beginning), our results provide hope that effective, practical remedies exist that can help people make better judgments and decisions. This insight is particularly important because “the people who have the greatest influence on the lives of others are likely to be optimistic and overconfident and to take more risks than they realize” (Kahneman, 2011, p. 256).

References

- Åstebro, T. (2003). The Return to Independent Invention: Evidence of Unrealistic Optimism, Risk Seeking or Skewness Loving? *The Economic Journal*, *113*(484), 226-239.
- Åstebro, T., Jeffrey, S., & Adomdza, G. (2007). Inventor perseverance after being told to quit: The role of cognitive biases. *Journal of Behavioral Decision Making*, *20*, 253-272. <http://dx.doi.org/10.1002/bdm.554>.
- Barber, B., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics*, *116*, 261-292. <http://dx.doi.org/10.1162/003355301556400>.
- Bechara, A., Damasio, H., & Damasio, A.R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*(3), 295-307.
- Bengtsson, C., Persson, M., & Willenhag, P. (2005). Gender and overconfidence. *Economics Letters*, *86*, 199-203. <http://dx.doi.org/10.1016/j.econlet.2004.07.012>.
- Bernardo, A. E., & Welch, I. (2001). On the evolution of overconfidence and entrepreneurs. *Journal of Economics & Management Strategy*, *10*(3), 301-330.
- Berner, E. S., & Graber, M.L. (2008). Overconfidence as a cause of diagnostic error in medicine. *American Journal of Medicine*, *121*(Supplement), S2-S23.
- Blavatsky, P. (2009). Betting on own knowledge: Experimental test of overconfidence. *Journal of Risk and Uncertainty*, *38*, 39-49.
- Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*, 129-148. <http://dx.doi.org/10.1037//0003-066X.36.2.129>.
- Budescu, D., Wallsten, T., & Au, W. (1997). On the importance of random error in the study of probability judgment. Part II: Applying the stochastic judgment model to detect systematic trends. *Journal of Behavioral Decision Making*, *10*, 173-188.

- Busenitz, L. W., & Barney, J. B. (1997). Differences between entrepreneurs and managers in large organizations: Biases and heuristics in strategic decision-making. *Journal of Business Venturing, 12*(1), 9-30.
- Camerer, C., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review, 89*, 306-318.
<http://dx.doi.org/10.1257/aer.89.1.306>.
- Cesarini, D., Johannesson, M., Lichtenstein, P., & Wallace, B. (2009). Heritability of overconfidence. *Journal of the European Economic Association, 7*, 617-627.
<http://dx.doi.org/10.1162/JEEA.2009.7.2-3.617>.
- Cesarini, D., Sandewall, Ö., & Johannesson, M. (2006). Confidence interval estimation tasks and the economics of overconfidence. *Journal of Economic Behavior and Organization, 61*, 453-470. <http://dx.doi.org/10.1016/j.jebo.2004.10.010>.
- Clore, G. L., & Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences, 11*, 393-399.
<http://dx.doi.org/10.1016/j.tics.2007.08.005>.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Colvin, C. R., & Block, J. (1994). Do positive illusions foster mental health? An examination of the Taylor and Brown formulation. *Psychological Bulletin, 116*(1), 3-20.
<http://dx.doi.org/10.1037/0033-2909.116.1.3>
- Colvin, R., Block, J., & Funder, D. (1995). Overly positive self-evaluations and personality: Negative implications for mental health. *Journal of Personality and Social Psychology, 68*, 1152-1162. <http://dx.doi.org/10.1037//0022-3514.68.6.1152>.
- Cooper, A. C., Woo, C. Y., & Dunkelberg, W. C. (1988). Entrepreneurs' perceived chances for success. *Journal of Business Venturing, 3*(2), 97-108.

- Cummins, R. A., & Nistico, H. (2002). Maintaining life satisfaction: The role of positive cognitive bias. *Journal of Happiness Studies*, 3(1), 37-69.
- Dohmen, T., Falk, A., Huffman, D., Schupp, J., Sunde, U., & Wagner, G. (2005). *Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey*. DIW discussion paper no. 1730. Retrieved from <http://ftp.iza.org/dp1730.pdf>.
- Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R. J. Sternberg & J. Davidson (Eds.), *The Nature of Insight* (pp. 365–395). Cambridge, MA: MIT Press.
- Dunning, D., & Story, A. (1991). Depression, realism, and the overconfidence effect: Are the sadder wiser when predicting future actions and events? *Journal of Personality and Social Psychology*, 61, 521-532. <http://dx.doi.org/10.1037//0022-3514.61.4.521>.
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5(3), 69-106.
- Ely, J. W., Graber, M. L., & Croskerry, P. (2011). Checklists to reduce diagnostic errors. *Academic Medicine*, 86, 307-313.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149-1160.
- Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology*, 3, 552-564.
- Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin*, 117, 39-66. <http://dx.doi.org/10.1037//0033-2909.117.1.39>.

- Freedman, J. L. (1978). *Happy People: What Happiness Is, Who Has It, and Why*. New York, NY: Harcourt Brace Jovanovich.
- Fu, T., Koutstaal, W., Fu, C., Poon, L., & Cleare, A. (2005). Depression, confidence, and decision: Evidence against depressive realism. *Journal of Psychopathology and Behavioral Assessment*, 27, 243-252. <http://dx.doi.org/10.1007/s10862-005-2404-x>.
- Galasso, A., & Simcoe, T.S. (2011). CEO overconfidence and innovation. *Management Science*, 57, 1469-1484.
- Gasper, K. (2004) Do you see what I see? Affect and visual information processing. *Cognition and Emotion*, 18, 405-421. <http://dx.doi.org/10.1080/02699930341000068>.
- Gasper, K., & Clore, G. L. (2000). Do you have to pay attention to your feelings to be influenced by them? *Personality and Social Psychology Bulletin*, 26, 698-711. <http://dx.doi.org/10.1177/0146167200268005>.
- Gerrards-Hesse, A., Spies, K., & Hesse, F. (1994). Experimental inductions of emotional states and their effectiveness: A review. *British Journal of Psychology*, 85, 55-78. <http://dx.doi.org/10.1111/j.2044-8295.1994.tb02508.x>.
- Goel, A. M., & Thakor, A. V. (2008). Overconfidence, CEO selection, and corporate governance. *The Journal of Finance*, 63(6), 2737-2784.
- Golec, J., & Tamarkin, M. (1995). Do bettors prefer long shots because they are risk-lovers, or are they just overconfident? *Journal of Risk and Uncertainty*, 11, 51-64. <http://dx.doi.org/10.1007/BF01132730>.
- Gosling, S., Rentfrow, P., & Swann, W. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37, 504-528. [http://dx.doi.org/10.1016/S0092-6566\(03\)00046-1](http://dx.doi.org/10.1016/S0092-6566(03)00046-1).
- Heath, C., Larrick, R.P., & Klayman, J. (1998). Cognitive repairs: How organizational practices can compensate for individual shortcomings. *Research in Organizational*

- Behavior*, 20, (pp. 1–37). Retrieved from <http://faculty-gsb.stanford.edu/heath/documents/rob-cognitive%20repairs.pdf>
- Hewig, J., Hagemann, D., Seifert, J., Gollwitzer, M., Naumann, E., & Bartussek, D. (2005). A revised film set for the induction of basic emotions. *Cognition and Emotion*, 19, 1095-1109.
- Hirshleifer, D., Low, A., & Theoh, S. H. (2012). Are overconfident CEOs better innovators? *Journal of Finance*, 67, 1457-1498.
- Hoelzl, E., & Rustichini, A. (2005). Overconfident: Do you put your money on it? *Economic Journal*, 115, 305-318. <http://dx.doi.org/10.1111/j.1468-0297.2005.00990.x>.
- Hoffrage, U. (2004). Overconfidence. In R. F. Pohl (Ed.), *Cognitive Illusions: Fallacies and Biases in Thinking, Judgement, and Memory* (pp. 235–254). Hove, UK: Psychology Press.
- Holt, C., & Laury, S. (2002). Risk aversion and incentive effects. *American Economic Review*, 92, 1644-1655.
- Holt, C., & Laury, S. (2005). Risk aversion and incentive effects: New data without order effects. *American Economic Review*, 95, 902-904.
- Ifcher, J., & Zarghamee, H. (2011). Positive affect and overconfidence: A laboratory investigation. *SCU Leavey School of Business Research Paper*, No. 11-02. Retrieved from <http://www.scu.edu/business/economics/research/upload/PAandOC-June2012.pdf>.
- Isen, A. M., Shalcker, T. E., Clark, M., & Karp, L. (1978). Affect, accessibility of material in memory, and behavior: a cognitive loop? *Journal of Personality and Social Psychology*, 36(1), 1-12.
- Isen, A. M., Daubman, K. A., & Gorgoglione, J. M. (1987). The influence of positive affect on cognition organization: Implications for education. In R. E. Snow & M. J. Farr

- (Eds.), *Connotative and affective process analysis* (pp. 125-154). Hillsdale, NJ: Lawrence Erlbaum.
- Isen, A. M., & Geva, N. (1987). The influence of positive affect on acceptable level of risk: The person with a large canoe has a large worry. *Organizational Behavior & Human Decision Processes*, 39(2), 145-154.
- Johnson, D. D. (2009). *Overconfidence and War: The Havoc and Glory of Positive Illusions*. Cambridge, MA: President and Fellows of Harvard College Corp.
- Kadous, K. (2001). Improving jurors' evaluations of auditors in negligence cases. *Contemporary Accounting Research*, 18, 425-444. <http://dx.doi.org/10.1092/GM8A-HNPH-LL3L-98FY>.
- Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, 39(1), 17-31.
- Kahneman, D. (2011). *Thinking, fast and slow*. London, UK: Penguin Books.
- Kirchler, E., & Maciejovsky, B. (2002). Simultaneous over- and underconfidence: Evidence from experimental asset markets. *Journal of Risk and Uncertainty*, 25, 65-85. <http://dx.doi.org/10.2139/ssrn.303387>.
- Koellinger, P. (2008). Why are some entrepreneurs more innovative than others? *Small Business Economics*, 31, 21-37.
- Koellinger, P., Minniti, M., & Schade, C. (2007). "I think I can, I think I can" – Overconfidence and entrepreneurial behavior. *Journal of Economic Psychology*, 28, 502-527. <http://dx.doi.org/10.1016/j.joep.2006.11.002>.
- Kuhn, G. J. (2002). Diagnostic errors. *Academic Emergency Medicine*, 9, 740-750.
- Kurtz, J. L., Wilson, T. D., & Gilbert, D. T. (2007). Quantity versus uncertainty: When winning one gift is better than winning two. *Journal of Experimental Social Psychology*, 43, 979-985.

- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology*, 4(863).
<http://dx.doi.org/10.3389/fpsyg.2013.00863>
- Larrick, R. P., Burson, K. A., & Soll, J. B. (2007). Social comparison and confidence: When thinking you're better than average predicts overconfidence (and when it does not). *Organizational Behavior and Human Decision Processes*, 102, 76-94.
<http://dx.doi.org/10.1016/j.obhdp.2006.10.002>
- Lerner, J. S., Gonzalez, R. M., Small, D. A., & Fischhoff, B. (2003). Effects of fear and anger on perceived risks of terrorism: A national field experiment. *Psychological Science*, 14, 144-150. <http://dx.doi.org/10.1111/1467-9280.01433>.
- Malmendier, U., & Tate, G. (2005). CEO overconfidence and corporate investment. *Journal of Finance*, 60, 2661-2700.
- Malmendier, U., & Tate, G. (2008). Who makes acquisitions? CEO overconfidence and the market's reaction. *Journal of Financial Economics*, 89, 20-73.
- Maxwell, S.E. (2004). The persistence of underpowered studies in psychological research: Causes, consequences, and remedies. *Psychological Methods*, 9, 147-163.
<http://dx.doi.org/10.1037/1082-989X.9.2.147>
- Merrow, E., Phillips K., & Myers, C. (1981). *Understanding Cost Growth and Performance Shortfalls in Pioneer Process Plants*. Santa Barbara, CA: Rand Corporation.
- Moore, D., & Healy, P. (2008). The trouble with overconfidence. *Psychological Review*, 115, 502-517. <http://dx.doi.org/10.1037/0033-295X.115.2.502>.
- Moore, D. A., & Small, D. A. (2007). Error and bias in comparative judgment: On being both better and worse than we think we are. *Journal of Personality and Social Psychology*, 92, 972-989. <http://dx.doi.org/10.1037/0022-3514.92.6.972>.

- Norman, G. R., & Eva, K. W. (2010). Diagnostic error and clinical reasoning. *Medical Education, 44*, 94-100.
- Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspectives on Psychological Science, 1*, 110-122. <http://dx.doi.org/10.1111/j.1745-6916.2006.00008.x>.
- Schaefer, P., Williams, C., Goodie, A., & Campbell, W. K. (2004). Overconfidence and the Big Five. *Journal of Research in Personality, 38*, 473-480.
<http://dx.doi.org/10.1016/j.jrp.2003.09.010>.
- Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A re-evaluation of the Life Orientation Test. *Journal of Personality and Social Psychology, 67*, 1063-1078.
- Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45*, 513-523. <http://dx.doi.org/10.1037//0022-3514.45.3.513>.
- Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy Scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35-37). Windsor, UK: NFER-Nelson.
- Seligman, M. (1991). *Learnt Optimism*. New York, NY: AA Knopf.
- Simonsohn, U. (2010). eBays crowded evenings: Competition neglect in market entry decisions. *Management Science, 56*, 1060-1073.
- Smith, V. (1982). Microeconomic systems as an experimental science. *American Economic Review, 72*, 923-955.
- Taylor, S., & Brown, J. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193-210.
<http://dx.doi.org/10.1037//0033-2909.103.2.193>.

- Taylor, S. E., & Brown, J. D. (1994). Positive illusions and well-being revisited: separating fact from fiction. *Psychological Bulletin*, *116*(1), 21-27.
- Taylor, S. E., Kemeny, M. E., Reed, G. M., Bower, J. E., & Gruenewald, T. L. (2000). Psychological resources, positive illusions, and health. *American Psychologist*, *55*(1), 99-109. <http://dx.doi.org/10.1037/0003-066X.55.1.99>
- Urbig, D., Stauf, J., & Weitzel, U. (2009). What is your level of overconfidence? A strictly incentive compatible measurement of absolute and relative overconfidence. *Discussion Paper Series 09-20*. Tjalling C. Koopmans Research Institute. Retrieved from <http://www.uu.nl/faculty/leg/NL/organisatie/departementen/departementeconomie/onderzoek/publicaties/DiscussionPapers/Documents/09-20.pdf>.
- Vautier, S., & Bonnefon, J.-F. (2008). Is the above-average effect measurable at all? The validity of the self-reported happiness minus others' perceived happiness construct. *Applied Psychological Measurement*, *32*(7), 575-584.
- Watson, D., & Clark, L. A. (1994). *The PANAS-X: Manual for the Positive and Negative Affect Schedule - Expanded Form*. Iowa City, IA: University of Iowa.
- Weinstein, N.D. (2003). Exploring the links between risk perceptions and preventive health behavior. In J. Suls & K. Wallston (Eds.), *Social Psychological Foundations of Health and Illness* (pp. 22–53). Malden, MA: Blackwell.
- Westermann, R., Spies, K., Stahl, G., & Hesse, F. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology*, *26*, 557-580. [http://dx.doi.org/10.1002/\(SICI\)1099-0992\(199607\)26:4<557::AID-EJSP769>3.0.CO;2-4](http://dx.doi.org/10.1002/(SICI)1099-0992(199607)26:4<557::AID-EJSP769>3.0.CO;2-4).
- Yeung, C. W. M., & Wyer Jr., R. S. (2004). Affect, appraisal, and consumer judgment. *Journal of Consumer Research*, *31*, 412-424. <http://dx.doi.org/10.1086/422119>.

Table 1

Effect sizes in prior studies

Source	Dependent variable	Control variables	β	$SE(\beta)$	t statistic	Cohen's d	N
Ifcher and Zargamhee (2011)	<i>OC</i>	No	0.94	0.81	1.16	0.22	107
		Yes	2.34	0.46	3.09	0.62	99
	<i>ROC</i>	No	1.04	0.88	1.12	0.22	104
		Yes	1.28	0.48	1.99	0.39	98
Pretest	<i>OC</i>	No	0.06	0.05	1.20	0.33	52
		Yes	0.05	0.05	1.00	0.28	52
	<i>ROC</i>	No	0.25	0.11	2.27	0.63	52
		Yes	0.16	0.07	2.29	0.63	52

Note: *OC* stands for absolute overconfidence, *ROC* for relative overconfidence. β is the unstandardized regression coefficient comparing the effect of the joy treatment group with that of the control group. β 's from Ifcher and Zargamhee are taken from Tables 4 and 5, models 1 and 3. Their control variables included sex, proportion of females in the session, demographic controls, disgust, embarrassment, and performance on a quiz (number of correct responses). The control variables in the analyses from our pretest included sex, age, age², financial stakes, study subject, and performance (percentage of correctly answered quiz questions). Excluding performance as a control variable yielded slightly higher estimated effects of the experimental treatment in our pretest. We do not know what the results of Ifcher and Zargamhee (2011) would look like if performance were excluded as a control variable. The t statistic is calculated as $\beta / SE(\beta)$. Cohen's d was calculated using t and N , following Lakens (2013). The standard errors of Cohen's d s in Ifcher and Zargamhee were 0.2; standard errors of Cohen's d s in our pretest were 0.29.

Table 2

Experimental Design

	Before experiment	Experiment			<i>N</i>
Control	-		O ₁	O ₂	51 (55)
Joy	X ₁		O ₁	O ₂	53 (58)
Joy awareness	X ₁	X ₂	O ₁	O ₂	55 (56)
Overconfidence		X ₃	O ₁	O ₂	54 (57)

Note. $N = 213$ (226). The numbers in brackets include participants who did not understand the experimental task, made unreasonable choices, or had computer problems. We excluded these 13 participants from further analyses, but our results are robust to the inclusion of these participants. X₁ = joyful affect induction (i.e., gift), X₂ = joy awareness treatment (i.e., scale), X₃ = overconfidence awareness treatment (i.e., text), O₁ = overconfidence task, O₂ = standard questions.

Table 3

Robustness Checks

		Absolute overconfidence					Relative overconfidence				
		Model 1	Model 2	Model 3	Model 4	Model 5	Model 1	Model 2	Model 3	Model 4	Model 5
Control	Joy treatment	$d = 0.36^*$	$d = 0.42^*$	$d = 0.37^\wedge$	$d = 0.53^{**}$	$d = 0.54^{**}$	$d = 0.29$	$d = 0.27$	$d = 0.23$	$d = 0.46^*$	$d = 0.54^{**}$
	Joy awareness	$d = 0.08$	$d = 0.02$	$d = 0.03$	$d = 0.09$	$d = 0.03$	$d = 0.13$	$d = 0.10$	$d = 0.11$	$d = 0.04$	$d = 0.02$
	OC awareness	$d = 0.18$	$d = 0.18$	$d = 0.17$	$d = 0.40^*$	$d = 0.31$	$d = 0.12$	$d = 0.09$	$d = 0.15$	$d = 0.06$	$d = 0.004$
Joy treatment	Joy awareness	$d = 0.45^{**}$	$d = 0.44^*$	$d = 0.41^*$	$d = 0.44^*$	$d = 0.52^{**}$	$d = 0.40^*$	$d = 0.37^\wedge$	$d = 0.33^\wedge$	$d = 0.42^*$	$d = 0.53^{**}$
	OC awareness	$d = 0.22$	$d = 0.24$	$d = 0.20$	$d = 0.12$	$d = 0.23$	$d = 0.40^*$	$d = 0.36^\wedge$	$d = 0.37^\wedge$	$d = 0.39^*$	$d = 0.54^{**}$
Joy awareness	OC awareness	$d = 0.27$	$d = 0.20$	$d = 0.20$	$d = 0.31$	$d = 0.28$	$d = 0.01$	$d = 0.005$	$d = 0.05$	$d = 0.02$	$d = 0.02$
Model diagnostics		$F(3, 209)$ = 2.38, $p = 0.07$	$F(3, 202)$ = 2.17, $p = 0.09$	$F(3, 195)$ = 1.78, $p = 0.15$	$F(3, 194)$ = 3.16, $p = 0.03$	$F(3, 208)$ = 3.51, $p = 0.01$	$F(3, 209)$ = 1.95, $p = 0.12$	$F(3, 202)$ = 1.59, $p = 0.19$	$F(3, 195)$ = 1.48, $p = 0.22$	$F(3, 194)$ = 2.38, $p = 0.07$	$F(3, 208)$ = 3.87, $p = 0.01$
partial η^2		0.03	0.03	0.03	0.05	0.05	0.03	0.02	0.02	0.04	0.05

Note. Cohen's d was calculated using M and SD following Lakens (2013). The standard errors of Cohen's d s are 0.2. Group differences with $p < 0.01$ are marked with **, $p < 0.05$ are marked with *; group differences with $p < 0.10$ are marked with \wedge . All analyses of variance were followed by an LSD post-hoc test.

Model 1: ANOVA for absolute and relative overconfidence between experimental groups.

Model 2: ANCOVA for absolute and relative overconfidence between experimental groups with demographic variables (age, age², gender, employment status) and order of choice sets as control variables.

Model 3: like Model 2 but additionally controlling for risk preferences, personality traits, general self-efficacy, and optimism.

Model 4: like Model 3 but additionally controlling for participants' performance as percentage of correctly answered quiz questions.

Model 5: ANCOVA for absolute and relative overconfidence between experimental groups with performance as a control variable.

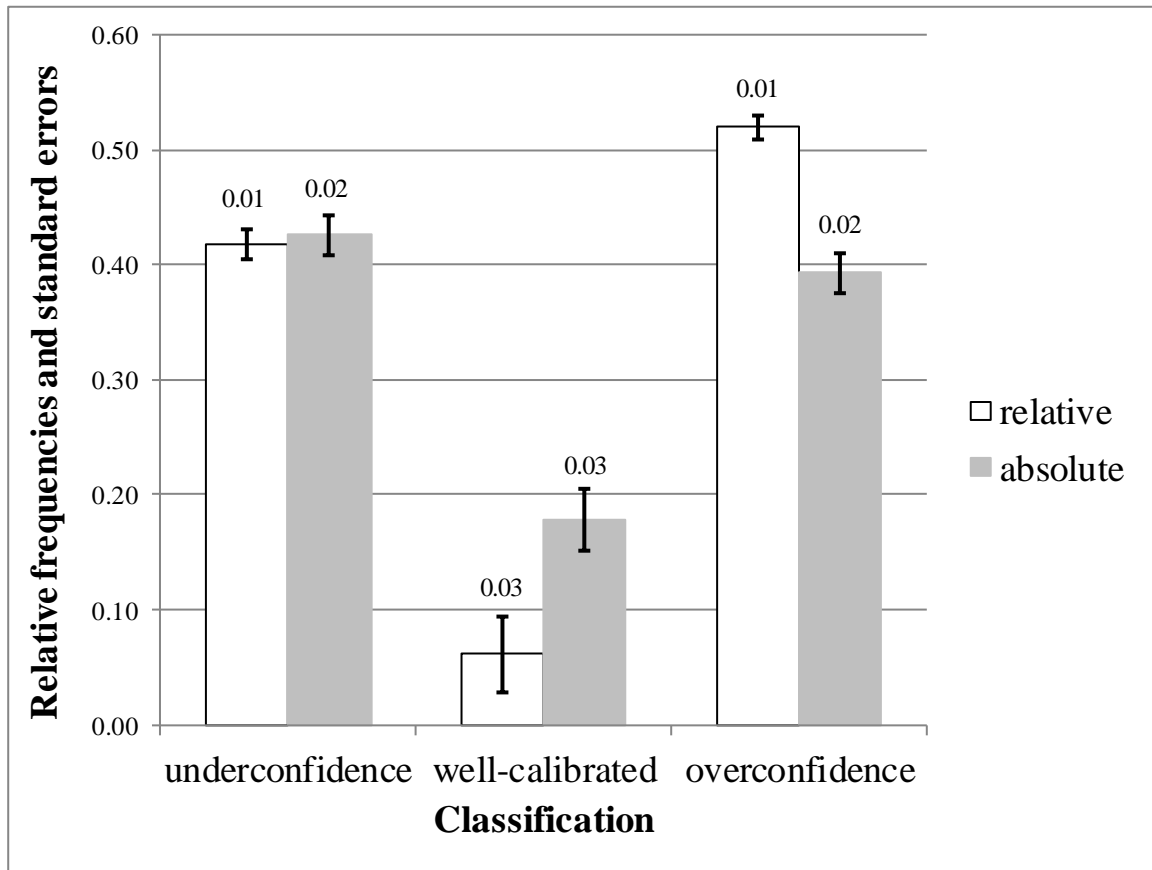


Figure 1. The relative frequency of under- and overconfident judgments for absolute (*ap*) and relative (*rp*) performance and their standard errors ($N = 213$).

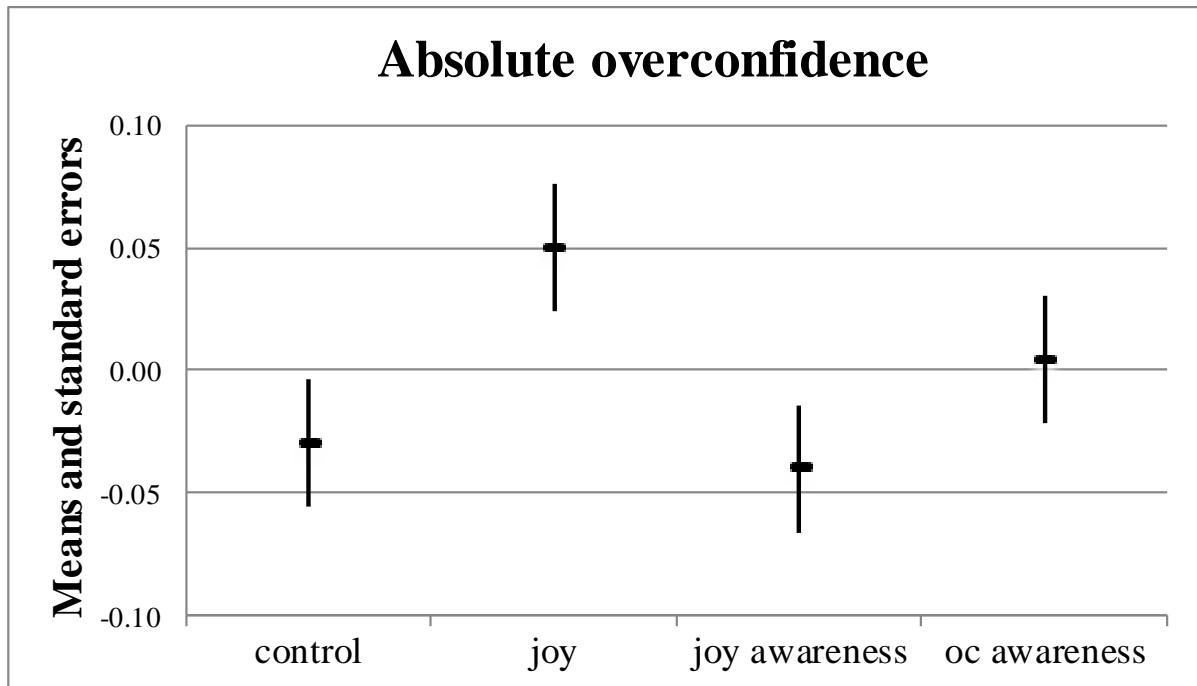


Figure 2. Absolute overconfidence across experimental groups, means, and standard errors ($N = 213$).

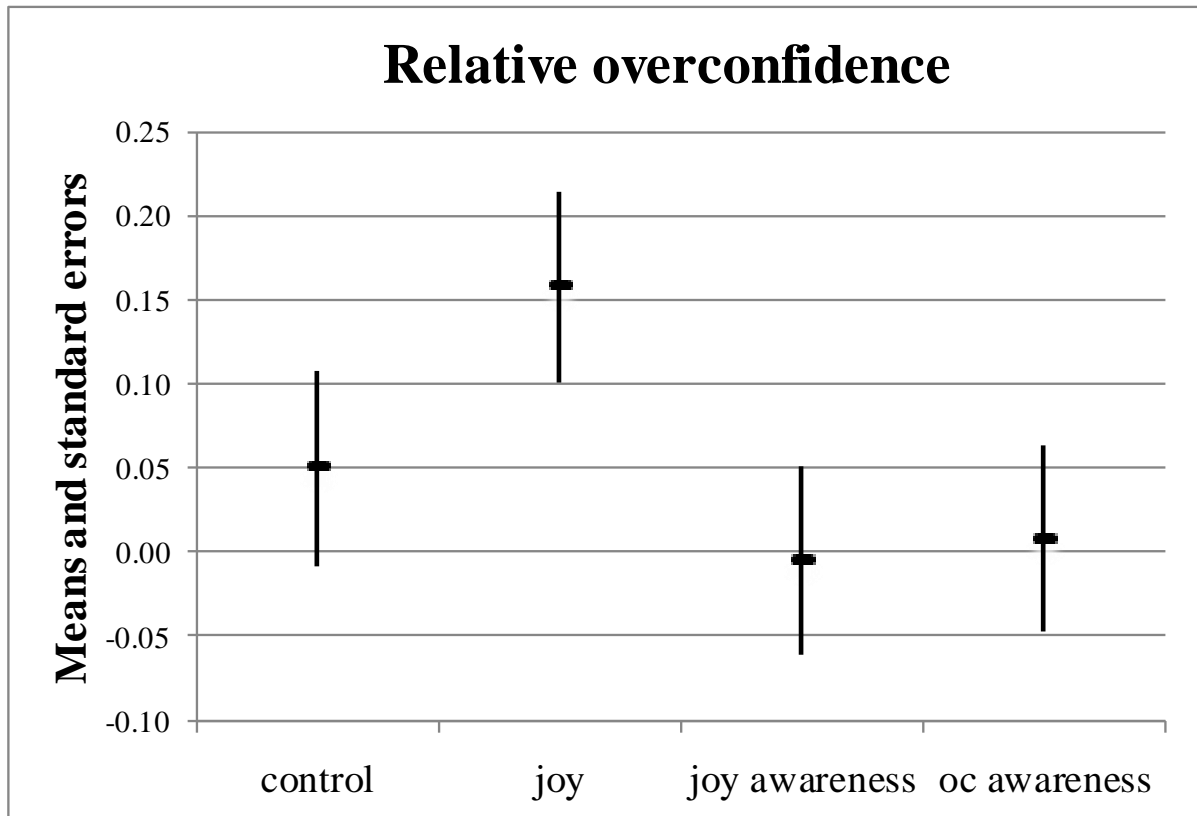


Figure 3. Relative overconfidence across experimental groups, means, and standard errors ($N = 213$).

APPENDIX 1: Pretest

The pretest was conducted at a large public university in Germany in 2010 with 52 students from various fields. The sample consisted of 22 males and 30 females with an average age of 25 years ($SE = 0.79$). The youngest and oldest participants were 18 and 45 years old. The pretest contained three different decision tasks, one of which was the current task on overconfidence.⁷ The participants completed the experiment with fixed (EUR 9, $N = 22$), low (up to EUR 11, $N = 24$), and high (possibility to win EUR 700, $N = 6$) financial stakes.

In each session, both a treatment (i.e., with a movie, comparable to the *joy awareness* group in our experiment, $N = 27$) and a control design (i.e., without a movie, comparable to the control group in our experiment, $N = 25$) were conducted with different participants. All participants completed 19 self-reported measures of their current affective state using three subscales (joy, fear and sadness) of the short version of the Positive and Negative Affect Schedule (PANAS-X; see Watson & Clark, 1994).

The treatment group watched a short, humorous movie clip (“When Harry met Sally” 1989, 3 min) (Hewig, Hagemann, Seifert, Gollwitzer, Naumann, & Bartussek, 2005) to induce positive affect (Gerrards-Hesse et al., 1994; Westermann et al., 1996). Because people are generally in a positive resting mood (e.g. Clore and Huntsinger 2007, Schwarz and Clore 1983) and positive affect is often more difficult to induce than negative affect (Westermann et al., 1996), we expected that the movie would induce only a small increase in joy among the participants in the treatment group. Prior to viewing the film clip, the participants were asked to become involved in the feelings suggested by the situation and to clear their minds of all thoughts, feelings and memories. Providing such explicit instructions increases the intensity of the affect induction (Westermann et al., 1996) and helps make

⁷ The other two tasks were unpublished measures of risk and ambiguity preferences.

participants aware of the irrelevant cause of their affective state. After the film clip, the participants were asked to complete a second affect measurement using the same PANAS-X subscales. All of the affect-related experimental elements in the treatment group very likely made the affect manipulation of the movie clip obvious to the participants. Hence, all participants were asked to reflect consciously upon their current affective state, but only the participants in the treatment group were confronted with a salient and non-relevant cause for their affect, i.e., the film clip.

The control group answered the PANAS-X subscales only once, at the beginning of the experiment. A mean comparison of self-reported joy using the relevant factor score from a Varimax-rotated principal components analysis on all 19 PANAS-X items⁸ shows that the control group was in a positive resting mood ($Mean = 0.13$, $SE = 0.22$) and that in the treatment group, the movie clip had the desired effect of increasing joy ($Mean = 0.87$, $SE = 0.24$), $t(50) = 2.23$, $p = 0.03$, $d = 0.63$.

Results

On average, the participants answered 4.83 ($SE = 0.23$) out of ten general knowledge questions correctly. The median performer answered five questions correctly. The best and the worst performances were eight questions ($N_8 = 4$) and one ($N_1 = 1$) question answered correctly, respectively. There was no statistical difference between the treatment ($Mean = 5$, $SE = 0.29$) and the control ($Mean = 4$, $SE = 0.36$) groups in their performance ($t(50) = -1.62$, $p = 0.11$, $d = 0.46$) or their range of performance (test of homogeneity of variances, Levene's $F(50) = 0.49$, $p = 0.48$, $d = 0.46$).

⁸ This procedure considers all available information, whereas average responses or factor scores, which are only based on the joy items, ignore mood differences among people arising from sadness or fear. However, all three possible measures of joy are almost perfectly correlated (*Pearson's* $r_s > 0.95$, $p < 0.01$).

A descriptive comparison of overconfidence across the experimental groups (Figure A1a and A1b) shows that the participants in the treatment group had better-calibrated judgments than the participants in the control group did. On average, judgments of absolute performance were exactly calibrated ($OC_{treatment} = 0.00$, $SE = 0.04$) in the treatment group. The treatment group was also better calibrated in terms of the participants' relative ability judgments ($roc_{treatment} = -0.06$, $SE = 0.08$) compared with the control group. The participants in the control group were on average overconfident about their absolute ($OC_{control} = 0.06$, $SE = 0.04$) and relative abilities ($roc_{control} = 0.20$, $SE = 0.06$). The differences between the treatment and the control group are statistically distinguishable for relative ($t[50] = 2.40$, $p = 0.02$, $d = 0.68$) but not for absolute ($t[50]=1.14$, $p = 0.26$, $d = 0.32$) overconfidence. Furthermore, the control group was on average strongly overconfident in a relative sense (one-sample t -test, $t_{mean=0}[24] = 3.20$, $p_{one-sided} = 0.002$, $d = 1.31$) and less strongly overconfident in an absolute sense (one-sample t -test, $t_{mean=0}[24] = 1.69$, $p_{one-sided} = 0.06$, $d = 0.69$).

Correlation results show that relative and absolute overconfidence tended to co-occur in our data (*Pearson's* $r = 0.57$, $p = 0.01$). Additional correlation results did not suggest any non-random distribution of relevant personal characteristics between the control group and the treatment group that could potentially bias the results.

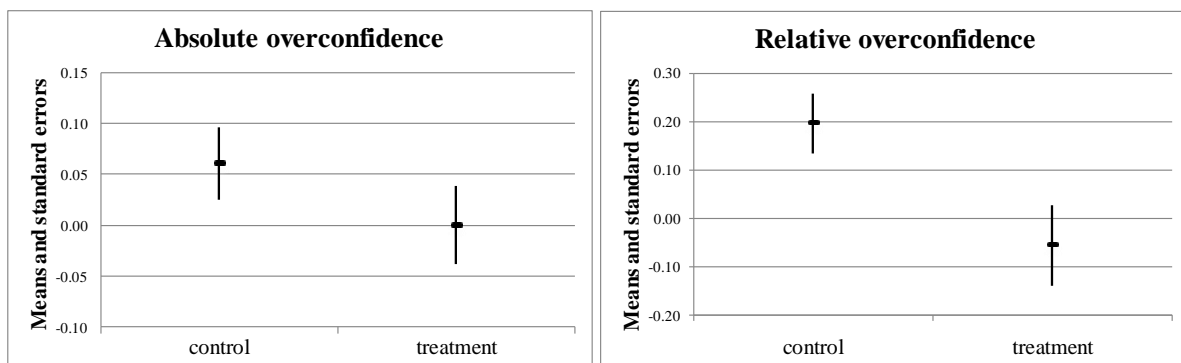


Figure A1. Absolute (left panel, A1a) and relative (right panel, A1b) overconfidence across experimental groups, means, and standard errors ($N = 52$).

From the pretest towards the main experiment

The results of our pretest suggest that the experimental manipulations we conducted had an effect on overconfidence. However, our experimental design in the pretest did not allow us to distinguish between two alternative interpretations of the results. The first explanation is that the higher level of joy in the treatment group actually improved judgment accuracy. This explanation would contradict affect-congruent findings and theories and the results of Ifcher and Zarghamee (2011), who independently conducted a relatively similar mood induction in their overconfidence experiment. The second explanation is that awareness of the irrelevant source of joy in the treatment group (i.e., watching a movie clip and completing the PANAS-X scale before and after watching the clip) led to the decrease in overconfidence, which would be consistent with the affect-as-information theory (Clore & Huntsinger, 2007; Schwarz & Clore, 1983). Importantly, the study of Ifcher and Zarghamee (2011) measured the self-reported moods of their participants *after* their overconfidence task, whereas our pretest measured moods *before* the task. Thus, our subjects were much more likely to realize before the overconfidence measurement that our purpose in showing them a movie clip was to manipulate their mood. This suggested to us that the awareness of the irrelevant source of joy and not the increased level of joy in the treatment group led to their improved judgment accuracy.

We designed our main experiment to parse these two competing explanations and to explicitly test the role of mood salience.

APPENDIX 3: Means (*M*), standard errors (*SE*), and correlation (*Pearson's r*) results (*N* = 213)

	<i>M</i>	<i>SE</i>	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Control	0.24	(-)									
(2) Joy	0.25	(-)	-0.32***								
(3) Joy awareness	0.26	(-)	-0.33***	-0.34***							
(4) OC awareness	0.25	(-)	-0.33***	-0.34***	-0.34***						
(5) OC	-0.01	0.01	-0.07	0.16**	-0.12^	0.03					
(6) ROC	0.05	0.03	0	0.16**	-0.08	-0.07	0.70***				
(7) Performance	0.46	0.01	-0.06	-0.02	0.05	0.03	-0.58***	-0.78***			
(8) Part	0.50	(-)	-0.10	0.05	-0.04	0.08	-0.02	-0.07	0.10		
(9) Age	21.00	0.16	-0.01	0.12^	-0.07	-0.04	0.10	0.14*	-0.01	0.01	
(10) Male	0.55	(-)	0	0	0.02	-0.01	-0.04	-0.11	0.17**	0.06	-0.03
(11) Employed	0.36	(-)	0.06	-0.13*	0.05	0.02	-0.02	-0.06	-0.05	-0.06	-0.22**
(12) Self-employed	0.04	(-)	-0.06	-0.07	-0.02	0.15*	-0.01	-0.07	0.11	0.07	0.12^
(13) Unemployed	0.24	(-)	-0.14*	0.08	0.06	-0.01	-0.07	0.05	-0.01	0.02	0.09
(14) Risk	4.34	0.10	-0.12^	0.11	-0.01	0.03	0.01	-0.02	0.06	-0.01	-0.03
(15) Extraversion	(-)	(-)	-0.04	0.10	0.03	-0.09	0.12^	0.06	-0.02	-0.01	-0.09
(16) Emot. Stab.	(-)	(-)	0.10	-0.07	0.05	-0.08	-0.03	-0.12^	0.10	0.02	-0.01
(17) Conscientious.	(-)	(-)	0.02	-0.06	0	0.03	-0.05	-0.01	-0.06	0.20**	0.08
(18) Openness	(-)	(-)	0.04	0.09	-0.02	-0.04	0.02	0.02	0.01	0.02	0
(19) GSE	3.06	0.03	0.13*	0.02	0.12^	-0.26***	-0.06	-0.09	0.16*	-0.08	-0.02
(20) Optimism	3.59	0.05	-0.03	0.06	0.02	-0.06	0.03	-0.10	0.11	-0.03	-0.09

	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
(11) Employed	0.03									
(12) Self-employed	0.10	0.14*								
(13) Unemployed	0.03	-0.33***	-0.12^							
(14) Risk	0.15*	-0.01	0.08	-0.04						
(15) Extraversion	-0.01	-0.01	0.01	0.04	0.29***	-				
(16) Emot. Stab.	0.23***	0.03	-0.04	-0.02	0.11^	-	-			
(17) Conscientious.	-0.11	0.05	0.09	-0.11	-0.05	-	-	-		
(18) Openness	0.02	0.04	0.08	-0.01	0.28***	-	-	-	-	
(19) GSE	0.16*	0.03	-0.01	-0.05	0.25***	0.17**	0.27***	0.20**	0.26***	
(20) Optimism	0.14*	0.11	-0.01	-0.12^	0.31***	0.35***	0.34***	0.04	0.08	0.28***

Note. Correlation coefficients are two-tailed, ^ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Joy awareness = joy awareness group, OC awareness = overconfidence awareness group, OC = absolute overconfidence, ROC = relative overconfidence, Performance = percentage of correctly answered quiz questions, Part = sequence of measures (1 = absolute overconfidence first), Risk = risk preferences, Emot. Stab. = emotional stability, Conscientious. = conscientiousness, GSE = general self-efficacy. Personality traits are based on factor scores.

ERIM Report Series <i>Research in Management</i>	
ERIM Report Series reference number	ERS-2012-001-STR
Date of publication	2012-01-03
Version	24-02-2014
Number of pages	46
Persistent URL for paper	http://hdl.handle.net/1765/30805
Email address corresponding author	T.Treffers@tue.nl
Address	Erasmus Research Institute of Management (ERIM) RSM Erasmus University / Erasmus School of Economics Erasmus University Rotterdam PO Box 1738 3000 DR Rotterdam, The Netherlands Phone: +31104081182 Fax: +31104089640 Email: info@erim.eur.nl Internet: http://www.erim.eur.nl
Availability	The ERIM Report Series is distributed through the following platforms: RePub, the EUR institutional repository Social Science Research Network (SSRN) Research Papers in Economics (RePEc)
Classifications	The electronic versions of the papers in the ERIM Report Series contain bibliographic metadata from the following classification systems: Library of Congress Classification (LCC) Journal of Economic Literature (JEL) ACM Computing Classification System Inspec Classification Scheme (ICS)