London Business School

LBS Research Online

J Benoit, J Dubra and G Romagnoli
Belief elicitation when more than money matters: controlling for "control"
Article

This version is available in the LBS Research Online repository: https://lbsresearch.london.edu/id/eprint/1570/

Benoit, J, Dubra, J and Romagnoli, G

(2022)

Belief elicitation when more than money matters: controlling for "control".

American Economic Journal: Microeconomics, 14 (3). pp. 837-888. ISSN 1945-7669

DOI: https://doi.org/10.1257/mic.20200017

American Economic Association
https://www.aeaweb.org/articles?id=10.1257%2Fmic.2...

Users may download and/or print one copy of any article(s) in LBS Research Online for purposes of research and/or private study. Further distribution of the material, or use for any commercial gain, is not permitted.

Belief Elicitation When More Than Money Matters: Controlling for “Control”.*

Jean-Pierre Benoît
London Business School

Juan Dubra
Universidad de Montevideo

Giorgia Romagnoli
University of Amsterdam

January, 2020

Abstract

Incentive compatible mechanisms for eliciting beliefs typically presume that the utilty of money is state independent or that money is the only argument in utility functions. However, subjects may have non-monetary objectives that confound these mechanisms. In particular, psychologists have argued that people favour bets where their ability is involved over equivalent random bets – a so-called preference for control. We propose a new belief elicitation method that mitigates the control preference. Using this method, we determine that under the ostensibly incentive compatible *matching probabilities method*, subjects report self-beliefs 18% higher than their true beliefs in order to increase control. Non-monetary objectives account for at least 68% of what would normally be measured as overconfidence. We also find that control manifests itself only as a desire for betting on doing well; betting on doing badly is perceived as a negative. Our mechanism can be used to yield better measurements of beliefs in contexts beyond the study of overconfidence.

*Contact: jpbenoit@london.edu, dubraj@um.edu.uy, G.Romagnoli@uva.nl.

We thank Ana Balsa, Alejandro Cid, Sebastian Galiani, Michael Jetter, Lewis Kornhauser, Madan Pillutla, Joël van der Weele and Peter Wakker for comments, and Facundo Danza, John Elder, Juan Pedro Gambetta and Bryan Kwan for outstanding research assistance. Juan Dubra acknowledges the financial support of ANII and Fondo Clemente Estable. The money for the experiments was provided by London Business School, New York University's Stern School of Business and CREED - University of Amsterdam.

Keywords: Elicitation, Overconfidence, Control. Experimental Methods.

Journal of Economic Literature Classification Numbers: D03; D83; C91

As economists have come to embrace the experimental paradigm long found in other disciplines, they have emphasized the benefits of incentivising subjects. Incentives both encourage subjects to participate in a meaningful way and guide experimenters in their interpretations of subjects' actions. Typical incentive protocols rely on monetary payments and the assumption that the utility for money is independent of the state, which routinely boils down to an assumption that money is the only argument in utility functions. Thus, an incentive compatible mechanism for eliciting beliefs is taken to be a mechanism in which subjects maximize their (state-independent) utility of money by truthfully reporting their beliefs.

However, while money is important, people also have non-monetary concerns. Researchers who ignore these concerns may end up with a distorted understanding of subjects' actions and beliefs. How large are possible distortions? We report on a new experimental design that allows us to obtain a measure of one type of distortion, which we summarize under the designation control, and to obtain a lower bound on the total non-monetary distortion. We find that distortions are sizable. When the matching probabilities method of Ducharme and Donnell (1973)[1] is used to elicit self-beliefs, subjects report beliefs 18% higher than their true beliefs out of control concerns. At least 68% of what would usually be interpreted as overconfidence is actually a willful overreporting.

Numerous experiments determine subjects' beliefs about themselves by presenting them with the opportunity to win a prize either based on their performance on a task or based on a random draw. In one format, subjects choose between a bet that yields the prize if their performance places them in, say, the top half of subjects and a bet that yields the prize with objective probability x (see, for example, Hoelzl and Rustichini (2005), Grieco and Hogarth (2009), Benoît, Dubra and Moore (2015), and Camerer and Lovallo (1999), which uses a similar format). The experimenter concludes that subjects who choose to bet on their performance believe they have a probability at least x of placing in the top half. In another format, subjects are asked to report the chances that they will place in the top half. The reports determine, in an incentive compatible manner, the probability that they will earn a prize based on their performance rather than from a random draw (see, for example,

[1]This method seems to have been invented by Smith (1961) and implemented by Ducharme and Donnell (1973), following on the Becker-Degroot-Marshack mechanism. It has been adapted by Grether (1981), Holt (2006), and Karni (2009). The literature does not use a consistent name for the method.

Hollard, Massoni and Vergnaud (2010), Andreoni and Sanchez (2014), Benoît, Dubra and Moore (2015), Bordalo, Coffman, Gennaioli and Shleifer (2019), Coutts (2019) and Möbius, Niederle, Niehaus and Rosenblat (2014)). The experimenter concludes that subjects who report the number y believe their chances of placing in the top half to be exactly y.

Yet, social scientists have identified (at least) two reasons the above conclusions about subjects' beliefs may overstate their actual beliefs.

1. **Betting on yourself: control.** People may have a preference for betting on themselves. Indeed, a long tradition in psychology holds that people have a desire for control in their lives, which may lead them to favour payments contingent on their performance over payments determined by probabilistically equivalent random devices.

2. **Positive statements.** People may derive benefits from making positive statements about themselves, either because they savour positive self-regard or to induce favourable treatment from others. This may lead them to exaggerate their odds of doing well on a task.

The presence of such non-monetary concerns is problematic for the experimenter. As Heath and Tversky (1991) write, "If willingness to bet on an uncertain event depends on more than the perceived likelihood of that event and the confidence in that estimate, it is exceedingly difficult – if not impossible – to derive underlying beliefs from preferences between bets." Heath and Tversky have in mind that subjects may choose to bet on their performance even if they believe the probabilities do not warrant it from a monetary perspective. For instance, a subject who thinks she has a 60% chance of placing in the top half of performers on a task may favour a bet on this eventuality over a bet with an objective 70% chance of paying off.

It is indeed difficult to disentangle subjects' beliefs from their disparate motivations by observing discrete choices they make. However, by comparing the choices subjects make under different conditions, we manage to isolate the desire for control and obtain a measure of the bias it introduces.

In our first experiment, self-beliefs are elicited using two different mechanisms. Under the first mechanism, subjects effectively choose between betting on themselves and betting on an objective random device. This mechanism employs the matching probabilities method, replicating previous literature. Under the second mechanism, subjects effectively choose between betting on themselves on one task and betting on themselves on another task. This novel design mitigates the control bias: no matter how subjects choose, they are betting on themselves. Both mechanisms are incentive compatible in money.

The implicit assumption in most of the existing literature is that the differences in the designs of the two mechanisms should not affect elicited beliefs. Since the two mechanisms are incentive compatible, when only money matters they should yield the same distribution of reports. Nevertheless, we find a significant discrepancy. With the matching probabilities mechanism that duplicates prior studies, subjects inflate their beliefs by 18% in order to shift weight towards bets on themselves (at the cost of reducing their overall chances of obtaining money). This experiment is run in the context of research on overconfidence and when our new mechanism is used to elicit beliefs we find that subjects display moderate overconfidence: the average reported chance of being in the top half is 54%, while beliefs should average to 50% in an unbiased population (the difference is statistically significant).

We run a second experiment in order to better understand control. Is it that people like to bet that they have done *well* on a task or that they like to bet on their performance, regardless of its quality? If the former, are people neutral about betting that they have done (unintentionally) poorly or do they actively dislike it and, if so, to what extent? These questions have received scant attention in the prior literature. In this second experiment, we address these questions by running a series of treatments in which subjects sometimes bet on doing well on a quiz and sometimes bet on having failed to do well.[2] We find that the control motivation manifests itself only as a desire for betting on doing well; a payment for doing badly is perceived as a negative, something we call an "anti-control" motive for bets on poor performance.

This refined understanding of the preference for betting on oneself is important for the analysis of the effect of control on elicitation methods that sometimes reward subjects for doing poorly, such as the state-of-the-art binarized scoring rule of Hossain and Okui (2013) and randomized scoring rule of Schlag and van der Weele (2013). In Section 2, we show exactly how control concerns affect the binarized scoring rule.

The elicitation technique we introduce rewards subjects for their performance on one of two tasks, rather than rewarding them either for their performance on a task or for the result of a random device. This design idea can be used independently of a desire to measure control and can be adapted to a variety of mechanisms, including the matching probabilities method, the binarized scoring rule, the randomized scoring rule, and the quadratic scoring

[2]Subjects are remunerated for correct quiz answers and they are not forewarned that they might later bet on a poor performance, so their incentive is to do well on the quiz.

rule (see surveys on incentive compatible elicitation by Schlag, Tremewan and van der Weele (2015) and Schotter and Trevino (2014)).

While our investigation is carried out within the overconfidence paradigm, its bearing is more general. Thus, Bordalo, Coffman, Gennaioli and Shleifer (2019), Cacault and Grieder (2019), Buser, Gerhards and van der Weele (2018), Sloof and von Siemens (2017), Massoni and Roux (2017), Coutts (2019) and Massoni, Gajdos and Vergnaud (2014) all use the matching probabilities method to elicit self-beliefs in studies which are not about overconfidence and, in each case, control is a confounder. Our study applies most readily to the elicitation of subjects' beliefs about themselves but, as Section 2.1 argues, it is useful beyond this realm.

In the economics literature, Owens, Grossman, and Fackler (2014) also investigates the implications of control for the interpretation of choices between bets. We discuss this paper in detail in sections 1.1 and 3.3. For now, we note that Owens et al., as well as Fehr, Herz and Wilkening (2013) and Bartling, Fehr and Herz (2014), presume that their elicitation methods are unaffected by control, whereas we measure the size of the control distortion in the elicitation methods.

1 Overstatement

In this section, we discuss some of the economics and psychology literature on non-monetary concerns that can lead subjects to misrepresent their beliefs.

1.1 Betting on Yourself: Control

Several studies conclude that people prefer bets on themselves to bets on probabilistically equivalent random devices.

In Goodie (2003), Goodie and Young (2007), and Heath and Tversky (1991, experiments 1, 2, and 3) subjects begin by answering a series of multiple choice questions and reporting the likelihoods that each answer is correct. They do not realize how these reports will be subsequently used.

Consider subjects who declare they have answered question i correctly with probability (about) p_i. In Goodie and in Goodie and Young, these subjects are split into two groups. In the first group, each subject chooses between (a) a bet that pays off if her answer to question i is correct and (b) the certainty-equivalent payment according to p_i. In the second group, each subject chooses between (a) a bet that pays off with an objective probability p_i and (b) the certainty-equivalent payment. Subjects in the first group choose the bet over the certainty-equivalent more often than subjects in the second group. In Heath and Tversky, each subject is given the choice between (a) a bet that pays off if her answer to question i is correct and (b) a bet that pays off with the objective probability p_i. Subjects take the first bet more often than the second bet, in domains in which they are competent.

These papers find that subjects' choices between betting on their answers and betting on a random device are not a simple reflection of the probabilities involved. Rather, subjects tend to display a bias towards betting on themselves–the more so, the more confident they are in their answers. Notice that when subjects choose to bet on themselves, they are choosing an ambiguous bet over an objective one. The interpretation is that the desire for control overcomes ambiguity aversion, at least when subjects have enough confidence in their answers. (Klein et al. (2010) explores the relation between ambiguity, controllability and competence).

Heath and Tversky argue that people have a special preference for betting on their answers in domains in which they are competent, while Goodie and Young dispute this interpretation and maintain that people have a general preference for control. As Goodie describes it, control is in play whenever the nature of the task is such that "a participant could take steps to favorably alter the success rate in subsequent administrations."[3] The exact reason a person might favour betting on herself – be it control, competence, or something else – is immaterial for our purposes and we, somewhat abusively, refer to any preference for betting on oneself as a *control* motivation.

While the findings of these papers are revealing, their methodologies do not permit a measurement of the value of control or the amount by which a preference for control would lead people to overstate their beliefs.[4] Moreover, the findings are weakened by the fact that subjects' reports of their likelihoods of correct answers are unincentivised. These papers

[3]Goodie talks of future administrations of the task as the subjects cannot better themselves in the current experiment and he wishes to distinguish control from the illusion of control. In the latter category, Li (2011) finds that subjects prefer a lottery in which they choose numbers over one in which the numbers are randomly selected, even though they recognize that the probabilities of winning is the same in the two. Our modelling accommodates both notions of control.

[4]Subjects' typically lost money by favouring bets on their answers – as much as 15% of earnings in one experiment in Heath and Tversky. It is impossible to tell to what extent these losses reflected overconfidence and to what extent a sacrifice for non-monetary objectives.

provide some motivation for our study but we do not undertake to match their frameworks.[5]

Owens et al. (2014) contrasts betting on one's own performance with betting on someone else's. Subjects are incentivised to report their beliefs that they will answer a question correctly and their beliefs that a randomly matched participant will answer a different question correctly. They are also asked to choose between a bet on their answer and a bet on the matched subject's answer. Based on the reported beliefs, if subjects care only about money they should choose to bet on themselves 56% of the time. Instead, subjects choose to bet on themselves 65% of the time, pointing to a preference for control. However, the interpretation of the results is somewhat clouded by the fact that the mechanism used for eliciting subjects' beliefs is itself prone to control issues. We return to this experiment in Section 3.3.

These four papers, and ours, can be viewed as exploring special cases of *source dependence* (Tversky and Wakker (1995)), whereby subjects consider the source of the uncertainty in addition to the probabilities involved. For axiomatizations that allow for source dependence see Klibanoff et al. (2005), Chew and Sagi (2008) and Gul and Pesendorfer (2015).

Karni and Safra (1995) and Almantier and Treich (2013) have shown that arbitrary state-dependent utility functions can pose insurmountable problems for truthful elicitation. Although some progress has been made in some cases (see, for example, Karni (1999)), we know of no prior attempt to develop an elicitation method that limits the distortionary impact of control.

1.2 Positive Statements: Self-Regard and Signalling

People like to say nice things about themselves, both out of self-regard and because sending out positive signals may induce favourable treatment from others. As Baumeister (1982) writes "The desire to be one's ideal self gives rise to motivations affecting both the private self and the public self ... It may also cause individuals to want an audience to perceive them as being the way they would like to be... The experimenter constitutes a real and important 'public' to the subject".

Burks et al. (2013) runs an experiment in which subjects take a quiz and are asked to predict the quintile into which they will place. Subjects also answer a personality traits questionnaire, which reveals that people with a high concern for social image tend to place themselves in high quintiles. The authors conclude that social signalling motives may lead

[5] Indeed, there are elements of these papers which we want to avoid. For instance, in Heath and Tversky's second and third experiments, subjects are asked to rate their knowledge of the subject matter, in addition to their probability of answering a question correctly, which could have an effect on their subsequent behavior.

subjects to overstate their beliefs. Ewers and Zimmermann (2015) asks subjects whether they believe their performance on a quiz was better or worse than the average performance of another group. Subjects' reports are either (a) only entered privately onto a computer screen or (b) entered onto a computer screen and also given orally in front of other subjects. The latter more public reporting results in significantly higher self-assessments. The authors conclude that subjects inflate their assessments in order to appear skillful to others.[6]

On the other hand, Benoît et. al. (2015) varies the perceived importance of a task that subjects undertake and, although a more important task should give subjects a greater motivation to appear competent to others, the variation produces no effect on reported placements. This may be because there are also costs to inflating one's self-assessment, including potentially looking or feeling foolish if actual performance falls short of predictions.

2 Formalism

We now incorporate the desire for control and for saying nice things into a model of utility. For ease of exposition, we develop our formalism in the context of the experiments we run, rather than setting out the most general formulation. Our simple model allows us to identify the effect of control in our experiments. In Section 3.2 we discuss conclusions that are independent of the specific modelling we adopt.

Consider an experiment where a subject undertakes a task for which her performance is described by a variable $\theta \in \{\theta^L, \theta^H\}$, where θ^L indicates a low, or poor, performance and θ^H indicates a high performance. The subject believes there is a chance μ that she will perform well, $\theta = \theta^H$, and she is asked for a report p of this belief. She might earn an amount of money m, depending on how well she performs, the number p she indicates, and random draws. If she has an initial wealth w and earns the amount m with (subjective) probability $r(p, \mu)$ and the amount 0 with probability $(1 - r)$, her expected monetary utility from the experiment is $ru(w + m) + (1 - r) u(w)$. We add two elements to this standard utility function:

1. **Control**. A subject derives an extra utility kick from money that is obtained for her performance, rather than through a random device: when she is paid m for achieving

[6]More precisely, Ewers and Zimmermann conclude that their findings are consistent with some people making reports that are higher than their actual beliefs and some having overconfident beliefs. Schwardmann and van der Weele (2019) find that people who can earn money by convincing others that they are high performers may also end up deceiving themselves.

performance level θ^i, $i = L, H$, she derives extra utility c_i beyond the utility of the money itself. To be precise, the subject earns extra utility c_i when she earns m and her performance θ is at level θ^i, but she would have earned 0 if instead her performance was at level $\theta^{j \neq i}$, ceteris paribus.[7] When the elicitation mechanism is such that the subject is paid for performance θ^i with probability $q_i(p, \mu)$, the expected utility kick is $c_i q_i$. A complex bet might involve the possibility of sometimes paying a subject for having done well, other times for having done poorly, so that in general the expected utility gain from control is $c_H q_H + c_L q_L$. Perhaps the most natural reading of the literature is that a subject derives a control benefit only from money obtained for having done well, not from money obtained for having done (unintentionally) poorly, so that $c_H > 0$ but $c_L \leq 0$. However, the literature is not very precise on this point and another possible reading is that $c_L > 0$. Experiment 1 examines the nature of c_H, while Experiment 2 also examines c_L.

2. **Self-regard and signalling**. A subject who believes $\theta = \theta^H$ with probability μ and reports p, gets an extra utility kick of $n(\mu)p$ from the report, where $n(\cdot) \geq 0$. If $n(\mu) \equiv 0$, then people derive no benefit from their reports per se. If $n'(\mu) < 0$, then higher types see less reason to inflate their reports. A more general formulation would give the kick as $x(\mu, p)$, with $x_2 \geq 0$. As self-regard/signalling motives are tangential to our study, we use the formulation $x(\mu, p) = n(\mu)p$ to simplify the analysis. We briefly discuss the more general formulation in Section 3.2. (This is a reduced form approach to incorporating self-regard and signalling benefits. See Burks et al. (2013) for a derivation of a signalling motive.)

A subject's total expected utility from participating in the experiment is

$$ru(w+m) + (1-r)u(w) + c_H q_H + c_L q_L + np.$$

Two important mechanisms. Under the matching probabilities mechanism, after a subject reports a belief p that $\theta = \theta^H$, a number x is drawn uniformly from $[0, 1]$. If $x \leq p$, the subject wins an amount m if her performance is high. If $x > p$, the subject wins m with probability x. Here, $r = p\mu + (1-p)\frac{p+1}{2}$, $q_H = p\mu$, and $q_L = 0$. Expected utility is maximized by a report $p^* = \mu + c_H\mu + n$. As this calculation indicates, the matching probabilities

[7] We safely omit any dependence of c_i on the amount of money m, as this amount does not vary within any of our experiments.

method is affected by control, and by c_H in particular. We explore this mechanism in detail in Experiment 1 below.

Under the binarized scoring rule, a random number z is drawn uniformly from $[0,1]$. The subject wins an amount m if and only if (a) $\theta = \theta^H$ and $z \geq (1-p)^2$ or (b) $\theta = \theta^L$ and $z \geq p^2$. Although the rule does not explicitly present subjects with a trade-off between winning based on their performance and winning based on a random draw, it does so implicitly, as we now show.

Suppose the subject reports $p \geq \frac{1}{2}$. If $z \geq p^2$, she wins m regardless of her performance; if $z < (1-p)^2$ she wins 0 regardless of her performance. In both cases, control plays no role. Control is at play when $(1-p)^2 \leq z < p^2$, as she then wins m if and only if she performs well. Here, $r = 1 - p^2 + \left(p^2 - (1-p)^2\right)\mu$, $q_H = \left(p^2 - (1-p)^2\right)\mu$ and $q_L = 0$. Expected utility from a report $p \geq \frac{1}{2}$ is maximized at $p^* = \mu + c_H\mu + \frac{n}{2}$.

Similar reasoning shows that if the subject reports $p < \frac{1}{2}$, control is at play when $p^2 \leq z < (1-p)^2$, as she then wins m if and only if $\theta = \theta^L$. Observe that she now earns money for a poor performance. Here, $r = 1 - (1-p)^2 + (1-\mu)\left((1-p)^2 - p^2\right)$, $q_H = 0$, and $q_L = (1-\mu)\left((1-p)^2 - p^2\right)$. Expected utility from a report $p < \frac{1}{2}$ is maximized at $p^* = \mu - c_L(1-\mu) + \frac{n}{2}$, when this is less than $\frac{1}{2}$. Although c_L does not matter for the matching probabilities method, it does matter for the binarized scoring rule. We analyze the nature of c_L in Experiment 2.

We note that the oft-used quadratic scoring rule is similarly subject to control distortions.

2.1 Beliefs About Objective Events

While we develop our analysis and run our experiments in the context of subjects' beliefs about themselves, control issues also arise in the elicitation of beliefs about external events where subjects are not directly involved. Indeed, Goodie, Goodie and Young, and Heath and Tversky develop their theory of control for such beliefs. Our model can apply here as well.

For instance, one way to elicit beliefs about the chances it will rain tomorrow is: i) ask subjects if they believe it will rain, ii) ask them for the probability that their prediction will prove correct, and iii) reward them using the matching probabilities method. Suppose a person answers that it will rain in step i). Let θ^H denote the event that it rains and θ^L the event that it does not rain. Our model now goes through exactly as before.

A more common elicitation procedure skips step i) and only asks subjects for the probability p of rain. The model applies, taking $p > \frac{1}{2}$ as a prediction that it will rain. That is, a person who answers there is, say, an 80% chance of rain tomorrow is presumed to prefer to be paid for "correctly predicting" rain, than based on a random device.

2.2 Further Considerations

Consider, for a moment, an experiment in which an individual is given a lottery ticket that pays m if she answers a question correctly and 0 otherwise. If her belief in her answer is μ then, factoring in control, the expected utility of the lottery is $\mu u(w+m) + (1-\mu)u(w) + \mu c_H$. The expected control benefit is μc_H, which is increasing in the subjective probability of a correct answer, when $c_H > 0$. Intuitively, a person who believes she has only a small chance of answering the question correctly, perceives little expected control benefit to being paid for a correct answer.

This feature of our modelling is consistent with experimental findings noted in Section 1.1 that subjects are more likely to exhibit a bias towards bets on their answers when they have a greater belief in the answers. We find mixed evidence of this effect in our data: in some cases, coefficients are of the correct sign, but in others they are not, and they are not always significant (in our pre-registration, we suspected there would not be enough power to establish the effect). We do not present the analysis here, but it is available online.

Technically, the difference between the two non-monetary elements, control and self-regard/signalling, as we have modelled them, is that the control benefit is contingent, only accruing when a subject is paid for her performance, while the self-regard/signalling benefit always accrues, by virtue of the subject's report. Our formalism can capture additional non-monetary motivations, as well as variations on the two we have considered. For instance, according to cognitive evaluation theory, a person's intrinsic motivation is higher when payment provides information about her competence level (see Ryan, Mims and Koestner (1983)). Hence, people respond more productively to rewards that are contingent on their good performance. An extra utility kick c_H for paid performance is one way of modelling this. As a variation on the self-regard benefit, it could be that statements made to an experimenter and statements made as inputs on a computer yield different benefits, so that $n(\mu)p$ is in fact the result of two different components.

3 First Experiment: Controlling for "Control"

Both our experiments were pre-registered. The pre-registration materials can be found in Appendix C and online at https://aspredicted.org/ip7te.pdf.

This experiment was run with subject pools from the CREED Lab at the University of Amsterdam and the MELESSA lab at Ludwig-Maximilians University in Munich, in the spring of 2020 using the software oTree (Chen, Schonger and Wickens, 2016).[8] Due to restrictions imposed by the Covid-19 pandemic, the experiment was conducted online. Arechar, Gächter and Molleman (2018) argue that data quality from online experiments is adequate and reliable. Moreover, the authors' biggest concern in M-Turk experiments is dropout and we had only 9 subjects drop out, for a 1.4% dropout rate.[9] The low dropout rate was possibly due to the fact that subjects were college students associated with the labs, the experimenter was available to answer questions via email, and payoffs were significant, averaging \$13.50. The experiment lasted approximately 45 minutes. Our sample consisted of 310 undergraduate students from the University of Amsterdam and 300 subjects from LMU.

The experiment comprises two treatments, which allows us to isolate and measure the control motive. The first treatment closely follows the matching probabilities method, widely used to elicit beliefs, notably in studies on overconfidence (for example, Möbius et al. (2014) and Benoît et al. (2015)).[10] With this design, beliefs are elicited by having subjects compare bets on their performance on a task with bets on a random device. The second treatment uses a new design in which beliefs are elicited by having subjects compare bets which all depend on their performance, on one of two tasks.

The main hypothesis is that there is a control motive to overstate placement in Treatment 1 but not in Treatment 2, while self-regard/signalling motives are the same in the two

[8] A very similar experiment, with the same scope, was run in the CREED Lab in 2018 with 313 in-person participants. However, an incorrect copy of the experimental code was used by mistake, causing subjects to receive incomplete or incorrect information about their performance. Given the key role played by this information, we considered the data to be unreliable. The new experimental design has been kept essentially the same and rerun with 610 subjects. (The results from the previous, unreliable, experiment were similar in character to our new results, although the effects were not as large.)

[9] Four subjects dropped out in the first part of the experiment due to connection problems and five subjects dropped out at the belief elicitation page, presumably because they could not pass the review questions.

[10] Our implementation of the matching probabilites method differs slightly from the usual format in that subjects win lottery tickets for cash rather than cash directly. This difference is in order to make Treatment 1 comparable to Treatment 2, where lottery tickets are needed to preserve risk neutrality.

treatments. As a result, the average reported placement should be higher in Treatment 1 than in Treatment 2, and the difference in average reports can be used to measure the control motive. Moreover, the distribution of reported beliefs in Treatment 1 should first order stochastically dominate the distribution of beliefs in Treatment 2.

Timeline of the experiment

The two treatments share the following timeline.

1. Subjects undertake a visual task in which, on 10 occasions, a string of numbers appears briefly on a computer screen, after which they are asked to reproduce the string. The difficulty of the exercise varies across repetitions in the length of the string and the duration for which the string is shown. All subjects see the same sequence of strings.

2. Call s_i the fraction, or share, of the ten repetitions of the task in which subject i correctly identifies the string. Each subject i is told s_i.

3. Subjects answer three sample questions, similar to questions they will later answer in a logic quiz. Before they answer the sample questions, they are informed of this similarity and of the fact that they will need to form an incentivised assessment of their projected quiz performance compared to others.

4. Subjects are told the median quiz score of people that took the same quiz on prior occasions. Each subject is asked to report the chance that she will place in the top half of previous quiz-takers. One of two (monetarily) incentive compatible methods, one for each treatment, is used to incentivise the reports. Details are given below.

5. Subjects take a logic quiz in which they answer twelve multiple choice questions. The subjects are ranked according to their scores, with ties broken randomly.

6. Subjects are paid based on their performances in the visual task and the quiz and their reported beliefs, in a way which depends on the treatment assignment and is elaborated upon below.

Two mechanisms for belief elicitation

The two treatments differ solely in the way in which the beliefs elicited in step 4 above are incentivised. The incentive mechanisms are summarized below. Details of the mechanisms, as well as the instructions provided in the experiment, are given in the Appendix.

Treatment 1. Suppose subject i has indicated a probability p_1 of placing in the top half. A number $x \in [0,1]$ is drawn uniformly. If $x \leq p_1$ the subject wins a lottery ticket that pays €20 with a 30% chance if her score is above the median score of previous experimental sessions, with ties broken randomly. If $x > p_1$, with probability x she wins the lottery ticket. In all other cases, she wins nothing.

Treatment 2. Suppose subject i has indicated a probability p_2 of placing in the top half. A number $x \in [0,1]$ is drawn uniformly. If $x \leq p_2$ the subject wins a lottery ticket that pays €20 with a 30% chance if her score is above the median score of previous experimental sessions, with ties broken randomly. If $x > p_2$, with probability x she wins a lottery ticket that pays €20 with probability M_i if she was successful in a randomly drawn instance of the visual task. In all other cases, she wins nothing. Note that here the subject's skill is at play even when $x > p_2$.

In Treatment 2, for subject i the probability the ticket pays off is $M_i = \frac{3}{10s_i}$, when $s_i \geq \frac{3}{10}$(recall that s_i is the fraction of correct answers on the visual task). When $s_i < \frac{3}{10}$, the probability M_i is capped at 1. Subjects are told the numerical value of M_i without being apprised of its dependence on s_i. Notice that when $s_i \geq \frac{3}{10}$, the conditional chance of winning the money from a draw of the visual task and the ticket paying off is $s_i \times \frac{3}{10s_i} = 30\%$.

To understand the incentive properties of the mechanisms in the two treatments, consider a subject i who estimates her chance of placing in the top half to be μ.

First suppose that she cares only about money. When $s_i \geq \frac{3}{10}$, she should truthfully report her subjective belief that she will place in the top half of subjects, regardless of the treatment in which she participates. To see this, think of the choice between i) a *placement bet* which yields €20 with probability $\frac{3}{10}$ if the subject places in the top half and ii) a *random bet* which, with probability x, yields €20 with probability$\frac{3}{10}$. The subject prefers the placement bet if $\mu > x$ and the random bet if $\mu < x$. The mechanisms in the two treatments implement this preference by effectively asking her for the threshold probability p that causes her choice to switch from the placement bet to the random bet. Clearly, she optimizes by declaring $p = \mu$.

When $s_i < \frac{3}{10}$, the mechanism in Treatment 2 is no longer (monetarily) incentive compatible. As indicated in the pre-registration, for both treatments we exclude the 12% of subjects for whom $s_i < \frac{3}{10}$ from the main empirical analysis. Accordingly, in the theoretical analysis we focus on the case $s_i \geq \frac{3}{10}$.

Now suppose that subject i also has non-monetary concerns. We first carry out an informal analysis not tied to our specific modelling.

- In Treatment 1, any utility she derives from making positive statements about herself gives her an incentive to exaggerate her reported belief p_1. On top of this, a declaration p_1 means that with probability p_1 winning the €20 is dependent on her performance on the quiz, while with probability $(1-p_1)$ winning depends completely on a random device. Any utility she derives from betting on herself gives her a further incentive to inflate her report, in order to shift weight onto earning money for doing well rather than for being lucky.

- In Treatment 2, as in Treatment 1, the subject may inflate her report in order to say nice things about herself. Now, however, she can only earn money when she has performed well, either on the quiz or on the visual task. Utility derived from betting on herself no longer gives a further incentive to distort.

- Because a preference for control provides an incentive to inflate in Treatment 1 but not in Treatment 2, we expect $p_1 > p_2$ when subjects have control motives. The difference in the reports, $p_1 - p_2$, can be used to establish measures of the control effect.

We now reason formally, adopting the normalizations $u(w) = 0$ and $u(w+20) = U$, where w is a subject's initial wealth.

In Treatment 1, a subject who believes she has a probability μ of placing in the top half and reports a probability p_1 has a subjective probability $p_1 \mu \frac{3}{10} + (1-p_1)\frac{(1+p_1)}{2}\frac{3}{10}$ of winning the €20. Note that $(1-p_1)$ is the chance that the random draw x is above p_1 and $\frac{(1+p_1)}{2}$ is then the average value of x. In addition to the potential money gain, the subject derives a control benefit c_H when she is paid for doing well on the quiz. The probability that she is paid for doing well – that is, the probability she earns money when she places in the top half but would not have earned it had she not placed in the top half – is $p_1 \mu \frac{3}{10}$. The subject also obtains a self-regard benefit $n(\mu) p_1$ from her report. She has a total expected utility of

$$\left(p_1 \mu + \frac{1-p_1^2}{2}\right)\frac{3}{10}U + p_1 \mu \frac{3}{10} c_H + n(\mu) p_1. \tag{1}$$

This is maximized by a report

$$p_1^* = \mu(1 + C_H) + N(\mu) \tag{2}$$

making the substitutions $N(\mu) = \frac{10}{3}\frac{n(\mu)}{U}$ and $C_H = \frac{c_H}{U}$.[11]

[11] More precisely, we should write $p_1^* = \min\{\mu(1+C_H) + N(\mu), 1\}$. About 6% of subjects across the two treatments declare a probability of 1.

If the subject cares only about money, so that $N(\mu) \equiv 0 = C_H$, then $p_1^* = \mu$. Hence, the mechanism is monetarily incentive compatible. If $N(\mu) > 0$ and/or $C_H > 0$, the subject overstates her beliefs. We can interpret μC_H as the subject's overstatement due to control concerns, $N(\mu)$ as the overstatement due to self-image concerns, and $\mu C_H + N(\mu)$ as the total distortion.

In Treatment 2, a subject who believes she has a probability μ of placing in the top half and reports a probability p_2 has a subjective probability $p_2 \mu \frac{3}{10} + (1 - p_2) \frac{(1+p_2)}{2} s_i \frac{3}{10 s_i}$ of winning €20 (for $s_i \geq \frac{3}{10}$). The probability that she earns the money for her performance, either on the quiz or on the visual task, is also $p_2 \mu \frac{3}{10} + (1 - p_2) \frac{(1+p_2)}{2} s_i \frac{3}{10 s_i}$. Her total expected utility is

$$\left(p_2 \mu + \frac{1 - p_2^2}{2} \right) \frac{3}{10} (U + c_H) + n(\mu) p_2. \tag{3}$$

This is maximized by a report

$$p_2^* = \mu + \frac{N(\mu)}{C_H + 1}, \tag{4}$$

making the same substitutions as above. (Note that we are assuming that the control benefit does not depend on the task involved. We relax this assumption in Section 3.2.)

If $N(\mu) \equiv 0$ then $p_2^* = \mu$, so the mechanism is monetarily incentive compatible. If $N(\mu) > 0$ then $p_2^* > \mu$ – a subject with self-regard/signalling objectives overstates. Note that a control motivation, $C_H > 0$, does not give a reason to overstate; on the contrary, it dampens the self-image inflation. The reason for this dampening is that the control incentive reinforces the impetus to report truthfully, since $p_2 = \mu$ maximizes both the probability that the subject earns money and the probability that she earns it for doing well (as doing well is the only way she can earn money).

The next proposition uses the difference in the two treatments to establish identification in the experiment. The proposition follows trivially from the observation that $p_1^* = p_2^* (1 + C_H)$.

Proposition 1 *For a given belief $\mu > 0$, the optimal reports satisfy $p_1^* > p_2^*$ if and only if $C_H > 0$.*

3.1 Identification

We adopt a between subject design, with each subject participating in either Treatment 1 or Treatment 2. The two groups are drawn from the same pool, hence we make the standard assumption that the expected values of their beliefs are the same – $E(\mu_1) = E(\mu_2) = E(\mu)$. To achieve identification (the statistical analysis follows in Section 5.1), we treat our samples as large, so that mean beliefs in the two groups are the same – $\bar{\mu}_1 = \bar{\mu}_2 = E(\mu)$ –, the sample average reported beliefs, $\bar{p_1}$ and $\bar{p_2}$, satisfy $\bar{p_1} = E_\mu(p_1^*) \equiv \bar{p_1}^*$ and $\bar{p_2} = E_\mu(p_2^*) \equiv \bar{p_2}^*$, and the mean value of N is the same in the two groups – $\bar{N}(\mu_1) = \bar{N}(\mu_2) = \bar{N}$.

Consider Treatment 1. The standard interpretation of results in this type of experiment is that a finding of $\bar{p}_1 > \frac{1}{2}$ indicates that the population is overconfident, since the mechanism is incentive compatible and the mean belief in a well-calibrated population should be $\frac{1}{2}$ (see Benoît and Dubra (2011)). However, using (2) and averaging, we obtain $\bar{p}_1^* = \bar{\mu} + \bar{\mu}C_H + \bar{N}$ and an alternative possibility is that $\bar{\mu} = \frac{1}{2}$ but $C_H > 0$ and/or $\bar{N} > 0$. Then $\bar{\mu}C_H$ is the mean overstatement due to control concerns, $\bar{N}$ is the mean overstatement due to self-image concerns, and $\bar{\mu}C_H + \bar{N}$ is the mean total distortion. It is impossible to tell on the basis of Treatment 1 alone to what extent, if any, a finding of $\bar{p}_1 > \frac{1}{2}$ reflects non-monetary concerns rather than overconfident self-evaluations.

Nevertheless, Treatment 1 and 2 can be combined to elucidate the role of non-monetary concerns. First, Proposition 1 yields a test for the sign of C_H. A significant difference in treatment averages, $\bar{p}_1 - \bar{p}_2 > 0$, implies that $C_H > 0$; that is, the desire for control distorts reported beliefs. Our experimental findings, discussed in greater statistical detail in Section 5.1, are that $\bar{p}_1 = 64.21\%$ and $\bar{p}_2 = 54.54\%$. The difference $\bar{p}_1 - \bar{p}_2 = 9.67$ is significant at the 1% level, confirming the hypothesis that $\bar{p}_1 > \bar{p}_2$. Moreover, the empirical distribution of p_1's first order stochastically dominates the distribution of p_2's, as predicted by the control hypothesis (see Figure 1 in section 5.1.)

We can leverage the model further. Using (2) and (4), and averaging within the groups, we obtain

$$\bar{\mu}C_H + \bar{N} = \bar{p}_1^* - \bar{p}_2^* + \frac{\bar{N}}{C_H + 1} \geq \bar{p}_1^* - \bar{p}_2^* = \bar{p}_1 - \bar{p}_2.$$

Thus, $\bar{p}_1 - \bar{p}_2$ gives a lower bound on the overstatement $\bar{\mu}C_H + \bar{N}$ in Treatment 1 that is due to non-monetary concerns rather than to overconfidence. Treatment 1 uses a standard-type incentive mechanism and finds that, on average, people report an overestimate of their chances of being in the top half of 14.21 percentage points. Of this, at least 9.67 percentage points come from a willful inflation rather than a miscalibration. Put differently, at least 68% of the measured overconfidence in this experiment comes from control and self-regard/signalling distortions.

We can be more specific about the control markup μC_H. Again using (2) and (4), we obtain

$$C_H = \frac{\bar{p}_1^* - \bar{p}_2^*}{\bar{p}_2^*} = \frac{\bar{p}_1 - \bar{p}_2}{\bar{p}_2} = 17.7\%. \tag{5}$$

On average, each subject in Treatment 1 inflates her report by about 18% to derive control benefit 0.18μ.

Recall that the marginal benefit of control is $c_H = C_H U$, where $U = u(w+20) - u(w)$, so that $c_H = 0.18\,(u(20+w) - u(w))$. In words, the marginal utility from inflating for control reasons is 18% of the added utility from a €20 gain.

The average report elicited from our second mechanism, which is designed to eliminate the control motive, is 54%. This average displays mild overconfidence, or a mild desire to say nice things, since 54% is statistically different from 50% (p value 0.0054 for the two-sided test).

3.2 Discussion of Modelling

Our mechanisms involve compound lotteries and there is some evidence that lab participants have difficulty evaluating such lotteries (see Starmer and Sugden (1991)). However, Harrison, Martinez-Correa, and Swarthout (2015) finds that when subjects are paid for each of their choices, rather than through the random lottery method, their behaviour is consistent with being able to reduce objective compound lotteries. Our study pays subjects for their choices, and the difference between treatments 1 and 2 only involves objective lotteries, so that our use of compound lotteries is less of a concern than it might otherwise be.

Even if we accept that there is less of a concern here, the mechanisms we use remain somewhat complicated.[12] Despite this, our formal analysis assumes that the mechanisms are perfectly understood. The extent to which subjects understand the mechanisms is an issue that confronts the elicitation literature in general.

We address this issue to some degree within the experiment by having subjects answer five multiple-choice questions about the mechanisms. In order to proceed, subjects must correctly answer all five questions. If they make a mistake, they are not told which answer(s) need to be rectified. Guidance by the experimenter was available (via email) and turned out to be necessary in only a handful of cases.

Still, it is worth stepping back for a moment to consider what conclusions obtain without assuming full comprehension by subjects, and, for that matter, without adopting our specific model.

In Treatment 2 a subject can only earn money for a successful performance, whereas in

[12] Arguably, the mechanism in Treatment 2 is more complicated than the mechanism in Treatment 1, but it is unclear what impact, if any, this might have on reports.

Treatment 1 a subject can earn money either for her performance or from a random draw. Even for a subject without a detailed understanding of the mechanisms, it should be apparent that control incentives are mitigated in Treatment 2. This mitigation leads to the prediction that the distribution of reported p_1's will first order stochastically dominate the distribution of p_2's, without any formal modelling. The confirmation of this prediction is good evidence for the existence of a control effect and for the effectiveness of our new elicitation design.

Our model permits sharper conclusions, at the cost of added assumptions. We assume that self-regard motives yield a benefit $n(\mu)p$, rather than using a more general formulation $x(\mu, p)$. The more general $x(\mu, p)$ yields similar results if the function is "well-behaved".[13] We have explored several alternate ways of modelling the control kick, again with similar results. Suppose a subject gets a kick from winning a lottery ticket based on her performance, rather than from winning money based on her performance. Put differently, she considers the lottery ticket itself to be the prize. Then, in Treatment 2 her expected utility is

$$\left(p_2\mu\frac{3}{10} + (1-p_2)\frac{(1+p_2)}{2}s_i\frac{3}{10s_i}\right)U + \left(p_2\mu + (1-p_2)\frac{(1+p_2)}{2}s_i\right)c + n(\mu)p$$

Making a parallel adaptation for Treatment 1, we still find that the total distortion in Treatment 1 is at least 9.67 percentage points. Two differences are that, with this modelling, Treatment 2 does not completely eliminate the control motive and p_2^* is now inversely related to s_i, instead of independent of it. In another formulation, a subject gets an added kick from the very act of betting on herself, rather than from a *successful* bet on herself; the total distortion remains at least 9.67.

The model assumes that money earned for success on the visual task and money earned for success on the quiz yield the same control benefit C_H. The average performance on the two tasks is not too dissimilar – 70% on the visual task and 60% on the sample quiz questions (which is all subjects saw before making their reports) – lending credence to the two tasks having similar control benefits, but we do not know if the differing natures of the tasks affect subjects dispositions (though we have no particular reason to think so).[14]

More critically, perhaps, the nature of payment for success on the visual task differs from the nature of payment for the quiz in two ways. First, payment for the visual task comes

[13] In particular, with the formulation $x(\mu, p)$ suppose that $x_2 \geq 0$, and $x_{22} \leq 0$. Writting $X = \frac{10}{3}\frac{x}{U}$, we now have $p_1^* = \mu(1 + C_H) + X_2(\mu, p_1^*), p_2^* = \mu + \frac{X_2(\mu, p_2^*)}{C_H + 1}$ and $\bar{\mu}C_H + \bar{X}_2(p_1^*, \bar{\mu}) = \bar{p}_1^* - \bar{p}_2^* + \frac{X_2(\bar{\mu}, p_2^*)}{C_H+1} \geq \bar{p}_1^* - \bar{p}_2^*$, which mirrors our previous analysis. Since the distribution of p_1s first order stochastically dominates that of p_2s, we have $C_H \geq \frac{\bar{p}_1 - \bar{p}_2}{\bar{p}_2}$.

[14] The two tasks were expressly constructed to be dissimilar, as we did not want performance on the visual task to yield (much) information about performance on the quiz.

from a random draw of the task. Second, the visual task has been completed before subjects report their beliefs. Arguably, these two elements could lessen the feeling of control derived from the visual task. (Although for Goodie (2003) the fact that a task has been completed does not offset the control motive, which depends upon participants being able to improve their performance in *subsequent* trials; see Section 1.1).

Let us now allow that the quiz and visual task may yield different benefits C_H^q and C_H^v, respectively, with $C_H^v \leq C_H^q$.

In Treatment 2, the probability that a subject who declares p_2 earns money for her performance on the quiz is $p_2\mu\frac{3}{10}$, while the probability that she earns money for her performance on the visual task is $(1-p_2)\frac{(1+p_2)}{2}\frac{3}{10}$. Her expected utility is

$$\left(p_2\mu + \frac{1-p_2^2}{2}\right)\frac{3}{10}U + \left(p_2\mu\frac{3}{10}C_H^q U + \frac{1-p_2^2}{2}\frac{3}{10}C_H^v U\right) + n(\mu)\,p_2,$$

which is maximized by a declaration

$$p_2^* = \mu\frac{1+C_H^q}{1+C_H^v} + \frac{N(\mu)}{1+C_H^v}. \tag{6}$$

Under this formulation, the control benefit from the visual task is

$$C_H^v = \frac{\bar{p}_1^* - \bar{p}_2^*}{\bar{p}_2^*} = \frac{\bar{p}_1 - \bar{p}_2}{\bar{p}_2} = 17.7\%.$$

Under the assumption $C_H^q = C_H^v = C_H$, we have $C_H = 17.7\%$ – the result given earlier as equation (5). Under the current assumption $C_H^v \leq C_H^q$, we have $C_H^q \geq 17.7\%$ so that the control benefit in Treatment 1 is greater than what we reported earlier.

The total non-monetary inflation in Treatment 1 is given by $\bar{\mu}C_H^q + \bar{N}$. We now have

$$\begin{aligned}
\bar{\mu}C_H^q + \bar{N} &= \bar{p}_1^* - \left(\bar{p}_2^* - \frac{N(\mu)}{1+C_H^v}\right)\frac{1+C_H^v}{1+C_H^q} \\
&= \bar{p}_1^* - \bar{p}_2^*\frac{1+C_H^v}{1+C_H^q} + \frac{N(\mu)}{1+C_H^q} \\
&\geq \bar{p}_1^* - \bar{p}_2^*
\end{aligned}$$

and 9.67% is again a lower bound on the inflation.

If anything, allowing $C_H^v \leq C_H^q$ strengthens our findings. On the other hand, if for some reason the control benefit is greater from a random draw of a visual task than from performance on the quiz – $C_H^v > C_H^q$ – our results, as stated, are weakened. Notice, however, that when $C_H^v > C_H^q$ we have $C_H^v = 17.7\%$ so that, in any case, control motives are significant.

3.3 Betting on Yourself or Someone Else

In Owens, Grossman, and Fackler (2014), subjects choose between a bet that pays \$20 if they answer a question correctly and a bet that pays \$20 if a matched subject answers a different question correctly. Let μ_s be a subject's belief that she will answer her question correctly and μ_m be her belief that her matched subject will answer his question correctly. The easiest behaviour to interpret is the use of a *cutoff* strategy. With this strategy, a subject bets on herself if $\mu_s - \mu_m > k$, for some number k. If $k = 0$, the subject maximizes her expected monetary payoff; if $k < 0$ the subject values control and is willing to sacrifice money in order to bet on herself; if $k > 0$ she prefers to bet on someone else. Owens et al. use the word *control* as an "umbrella term" that encompasses any reason a person might favour a bet on herself. This includes choosing to bet on yourself to send a positive signal.

The beliefs μ_s, μ_m are not known to the experimenters. Rather, subjects are incentivised to report them by way of a matching probabilities method similar to the one we use in Treatment 1. Individual behaviour is evaluated with respect to the (observable) reports, p_s and p_m. A person is deemed to follow a cutoff strategy if she bets on herself when $p_s - p_m > k$, for some k. The authors determine that the behaviour of 82% of the subjects is consistent with a cutoff strategy. When $k < 0$, a subject is said to exhibit a preference for control.

Let us apply our modelling to this experiment. To begin, we keep things simple and assume that a) subjects have only a pure control motive, so that $c_H > 0$ but $n(\cdot) \equiv 0$, and b) they evaluate money won for someone else's performance purely in monetary terms. Under these assumptions, the elicited beliefs are given by

$$p_s^* = \mu_s (1 + C_H) \text{ and } p_m^* = \mu_m, \tag{7}$$

using the normalizations $u(w) = 0, u(w + 20) = U$, and $C_H = c_H/U$.

Now consider a subject's decision whether to bet on herself or bet on her match. Using our modelling, her payoff for betting on herself is

$$\mu_s u(w + 20) + (1 - \mu_s) u(w) + \mu_s c_H = \mu_s U + \mu_s C_H U, \tag{8}$$

while the payoff for betting on her match is

$$\mu_m u(w + 20) + (1 - \mu_m) u(w) = \mu_m U. \tag{9}$$

A subject chooses to bet on herself if $\mu_s - \mu_m > -\mu_s C_H$. If $c_H > 0$, as we find on average, then the unobservable cutoff $k = -\mu_s C_H$ is negative.

In terms of observables, from (7) we have that $\mu_s - \mu_m > -\mu_s \frac{c_H}{U}$ if and only if $p_s - q_s > 0$. Although the true cutoff k is negative, the measured cutoff $\hat{k}$ should be zero. Put differently, $\hat{k} = 0$ even for a subject with a positive control motivation (or a negative one, for that matter). In line with this reasoning, in one of their analyses, Owens et al. determine that, of the subjects with a cutoff behavior, 65% have a behavior that is consistent with a cutoff of 0. When these subjects are counted as not having a control motivation, our analysis implies that control is under-measured. In their conclusion, Owens et al. also reason that they have found a lower bound on the effect of control incentives.

Although the above reasoning suggests that the measured cutoff should be 0, in fact 26% of subjects display a strictly negative cutoff (and 9%, a strictly positive cutoff). This discrepancy can be reconciled with our modelling in several ways.

1. When given a direct choice between a bet on themselves and a bet on another person, some subjects may feel an extra push to bet on themselves. This push could be because of the positive signal sent by betting on oneself over someone else, because of the inherently larger ambiguity in a bet on someone else, or for some other reason. Such a push is consistent with the discussion in Owens et al. of the various reasons subjects may prefer bets on themselves. In terms of the above analysis, the simplifying assumptions a) and b) may not both hold.

2. The incentive to inflate for control may be especially salient to subjects in this experiment, where subjects are presented with a direct choice between two bets in contrast to the more elaborate matching probabilities mechanism.

3. Procedural details in this experiment and in ours may (inadvertently) play a role in the results.

The distinction between i) self-bets versus bets on someone else and ii) self-bets versus bets on a random device is an interesting one that our experiment and theory does not explore.

To directly test for the presence of control effects in the Owens et al. setup, we pre-registered and replicated their experimental procedure. Immediately following our previously described experiment, subjects were shown a multiple choice question selected from the Owens et al. study, together with possible answers, for 15 seconds. Subjects then reported their belief that they would later correctly answer the question when given forty-five seconds. This procedure was repeated four more times. For subjects in Treatment 1, reports were elicited using a matching probabilities method, as was done in Owens et al. For subjects in Treatment 2, beliefs were elicited with our new mechanism.

Our variable of interest is the average reported probability of answering the questions correctly. For each individual, we average her five answers and then average across subjects. As we describe in Section 5.1, the effect of Treatment 1 is to increase the average reported chance by 3 percentage points, a difference which is significant. Thus, the elicitation method used in Owens et al. is vulnerable to control distortions. Our findings suggest these distortions might be relatively small in this setup, although the subjects may have been tiring as this experiment immediately followed the previous one.

These findings are not directly comparable to the findings of our main experiment, as subjects were not asked about their relative placement here, the prize money was different and the money was disbursed using the random lottery method.

4 Second Experiment: The Meaning of Control

Experiment 2 was run in person in fall 2016, at the CREED laboratory of the University of Amsterdam. One hundred ninety-six undergraduates participated. The average payment for the experiment was €17.6 euro and the average duration was 50 minutes. There was no overlap in the samples of the two experiments.

The experiment seeks a better understanding of the control motivation. Our first experiment showed that people have a positive bias for bets that pay off when they do well. But how do they feel about bets that pay for an (unintentional) poor performance? Do these bets also yield a positive control benefit or are they undesirable in this regard? The answers are important for a proper understanding of the control motivation and for the analysis of incentive mechanisms that sometimes reward poor performance, such as the binarized scoring rule.

4.1 Three Treatments

Experiment 2 involves three treatments (the appendix provides the instructions that were used). In all three treatments subjects first take a quiz consisting of twenty multiple-choice questions. With a 50% chance they are paid €0.50 for each correct answer; with a 50% chance they are paid according to one of the mechanisms described below. (When taking the quiz, subjects are not aware of the nature of the mechanisms to follow, so that, presumably, their incentive is to do well on the quiz).

Treatment 1

A subject reports whether she prefers to bet she placed in the *top* half on the quiz or to bet on a random device for each objective chance 0%, 2%, ... ,100%. An even integer $x \in [0, 100]$ is drawn uniformly. If the subject indicated she prefers the placement bet for x, she wins €10 if she placed in the *top* half on the quiz, with ties broken randomly. If she indicated she prefers the random bet, she wins €10 with chance x. In all other cases, she wins nothing.

Treatment 2

A subject reports whether she prefers to bet she placed in the *bottom* half on the quiz or to bet on a random device for each objective chance 0%, 2%, ... ,100%. An even integer $x \in [0, 100]$ is drawn uniformly. If the subject indicated she prefers the placement bet for x, she wins €10 if she placed in the *bottom* half on the quiz. If she indicated she prefers the random bet, she wins €10 with chance x. In all other cases, she wins nothing.

Treatment 3

This treatment is a mixture of the first two.

A subject reports whether she prefers to bet she placed in the *top* half on the quiz or to bet on a random device for each objective chance 0%, 2%, ... ,100%. A coin is flipped and an even integer x is drawn. Suppose the coin comes up heads. If the subject indicated she prefers the placement bet for x, she wins €10 if she placed in the *top* half on the quiz. If she indicated she prefers the random bet, she wins €10 with chance x. Suppose the coin comes up tails. If the subject indicated she prefers the placement bet for $100 - x$, she wins €10 if she placed in the *bottom* half on the quiz. If she indicated she prefers the random bet, she wins €10 with chance $100 - x$. In all other cases, she wins nothing.[15]

Let p_1 be the highest probability for which a subject reports she prefers to bet on herself to the random device in Treatment 1; let q_2 and p_3 be the highest probabilities in Treatment 2 and 3. The above procedures implement the probability matching method where we interpret p_1 and p_3 to be a subject's reported belief she will place in the top half, and q_2 her reported belief she will place in the bottom half. (This interpretation is not possible for the 4 subjects out of 196 who made (irrational) non-monotonic choices.)

[15] In actuality, for half of the subjects in this treatment, the question was framed as a bet on placing in the bottom half, rather than in the upper half. To both groups it was explained that, depending on the results of the toss of the coin flip, they would end up betting either on their placement in the upper half or in the lower half. We found no difference between the two frames of choice (p-value = 0.677)

On a conceptual level, Treatment 1 here mimics Treatment 1 in the first experiment.[16] Subjects have an incentive to inflate their implicit reports, both for self-regard/signalling reasons and in order to bet on themselves doing well.

Treatment 2 has no parallel in Experiment 1. While self-regard/signalling concerns operate exactly as in Treatment 1 – subjects have an incentive to implicitly underreport the probability of placing in the bottom half, which is equivalent to overreporting the chance they end up in the top half –, control considerations are different. Here, subjects can be rewarded for doing poorly but not for doing well. In terms of our formalism, the parameter c_L, rather than c_H, now plays a role.

4.2 Reporting Incentives

We first analyze reporting incentives, making the substitution $q_2 = 1 - p_2$ and adopting the normalizations $u(w+0) = 0$ and $u(w+10) = 1$, where w is a subject's initial wealth.

Consider a subject who estimates her chance of placing in the top half to be μ and reports this chance as p_1, if in Treatment 1; effectively reports it as $p_2 = 1 - q_2$, if in Treatment 2; and reports it as p_3, if in Treatment 3.

In Treatment 1, she has an expected utility of

$$p_1 \mu + \frac{1 - p_1^2}{2} + c_H p_1 \mu + n(\mu) p_1,$$

which is maximized at

$$p_1^* = \mu (1 + c_H) + n(\mu). \tag{10}$$

In Treatment 2, she has an expected utility of

$$(1 - p_2)(1 - \mu) + \frac{2p_2 - p_2^2}{2} + c_L (1 - p_2)(1 - \mu) + n(\mu) p_2,$$

which is maximized at

$$p_2^* = \mu - c_L (1 - \mu) + n(\mu). \tag{11}$$

In Treatment 3, she has an expected utility of

$$\frac{1}{2}\left(p_3 \mu + \frac{1 - p_3^2}{2} + c_H p_3 \mu\right) + \frac{1}{2}\left((1 - p_3)(1 - \mu) + \frac{2p_3 - p_3^2}{2} + c_L (1 - p_3)(1 - \mu)\right) + n(\mu) p_3,$$

[16] In contrast to Experiment 1, here subjects make their predictions after having taken the test rather than after having seen sample questions, since they will sometimes bet on doing poorly. Because of this and other differences, the beliefs elicited in the two experiments are not directly comparable. This has no consequences for our analysis.

which is maximized at

$$p_3^* = \mu + \frac{1}{2} c_H \mu - \frac{1}{2} c_L (1 - \mu) + n(\mu). \tag{12}$$

We exploit these expressions in the next section.

4.3 Identification

We again analyze mean behaviour across treatments. From (10), (11), and (12), the theory demands that the optimal choices satisfy $\overline{p}_3^* = \frac{1}{2}\overline{p}_1^* + \frac{1}{2}\overline{p}_2^*$. Thus, Treatment 3 does not add anything to the estimation of the parameters but serves as a consistency check of the theory. The theory receives confirmation – or, at least, is not rejected – as we find that $\overline{p}_1 = 66.2\%$, $\overline{p}_2 = 67.9\%$ and $\overline{p}_3 = 66.7\%$ and, as we show later, we cannot reject $p_1^* = p_2^* = p_3^*$.

Given $\overline{p}_1^* = \overline{p}_2^*$, (10) and (11) together imply that

$$c_L = -c_H \frac{\mu}{1 - \mu}. \tag{13}$$

Experiment 1 established a strictly positive, and statistically significant desire for betting on one's success. The results of this experiment indicate an anti-control motive on money won for poor performances. This finding is consistent with Heath and Tversky's (1991) finding that subjects favour an ostensibly fair random bet over a bet that pays when they have answered a question *incorrectly*. Our result goes further, indicating that the utility loss from a payment for doing poorly, $c_L(1 - \mu)$, is the exact negative of the utility gain from a payment for doing well, $c_H \mu$.

On its own, the result $p_1^* = p_2^* = p_3^*$ allows for many interpretations, including that there is no control motivation and that elicited beliefs are independent of the mechanism used. We rely on the results of Experiment 1, as well as the results of other experiments that have determined control is a factor, for our interpretation. The conclusion that being paid for doing badly yields the negative of being paid for doing well is fairly intuitive.

As we noted in Section 2, c_L plays a role in the binarized scoring rule. We can now see that under this rule, control objectives lead a subject with belief μ to report $p^* = \mu + c_H \mu + \frac{n}{2} = \mu - c_L(1 - \mu) + \frac{n}{2}$. The mechanism we introduced to eliminate control distortions under the matching probabilities method can be adapted to eliminate control distortions with this rule too.

5 Statistical Analysis of the Experiments

In this section, we provide a statistical analysis of the results.

5.1 Experiment 1

Six hundred and ten undergraduates participated and completed the experiment and were randomly assigned to either Treatment 1 (N=306) or Treatment 2 (N=304). The randomization was successful in ensuring a good gender balance, with 59% females in Treatment 1 and 54% females in Treatment 2. The randomization was also balanced in terms of performance in the sample questions, a predictor of both placement and actual performance in the subsequent test (the mean number of correct sample questions was 1.82 and 1.77 in Treatment 1 and Treatment 2 respectively; difference not significant).

In accordance with our pre-registration, we exclude the 74 subjects with a success rate below 30% in the visual task from the main analysis, as the mechanism in Treatment 2 is not incentive compatible for them. We exclude these subjects for both treatments to avoid introducing a selection effect (results are unchanged if we include them in Treatment 1).

The main hypothesis is the existence of control motives to overstate beliefs in Treatment 1 but not in Treatment 2, while self-regard/signalling motives are the same in the two treatments. Formally, as pre-registered, we run two tests. First, we test if the average placement $\bar{p}_1$ in Treatment 1 is statistically larger than the placement $\bar{p}_2$ in Treatment 2 by running a regression of individuals' p_i on a constant and a dummy indicating Treatment 1, and checking for the significance on the dummy-coefficient with an independent two-sample one-sided t-test, which is appropriate given the sample size (N=536). The constant in the regression (Model A in Table 1) is 54.54 which is by design the average response in Treatment 2. The coefficient on the Treatment 1 dummy is 9.67; this measures how much larger the average response in T1 is compared to T2. The *t-test* supports the hypothesis ($t = 4.8$ and p-value $\simeq 0.0000$, $N = 536$).

In both treatments, placement might depend on variables such as the subject's gender, the score on the sample questions, or the percent probability M of earning the €20 in the lottery (in Treatment 2). Thus, our second pre-registered test is a regression where we add those variables (and dummies for lab, or major of the subjects). Major dummies are as follows: Economics, Economics and Business and related [ECON]; Psychology Politics Law and Economics (a selective interdisciplinary curriculum at the University of Amsterdam) [PPLE]; Other social sciences and Humanities; Science, Technology, Engineering and Mathematics and related [STEM]; Other [baseline dummy]. These dummies are not significant, but the order of confidence indicated by each major seems reasonable, with STEM having greater confidence in doing well in a logic quiz, followed by Economics, and then by Other Social Sciences and PPLE.

Among the controls is *Sample Score*, i.e., the number of correct answers given in the three sample questions. This variable is a signal that subjects can use to infer how well they will perform in the quiz (which, they are told, is based on questions similar to the sample questions).[17] As expected, a better performance on the sample quiz significantly increases the reported placement probability. Gender, included as "Male" or "Other" (Female omitted), has no significant effect, although the coefficient for males is positive.

Table 1 includes two non-preregistered robustness tests. We note that in both these checks, the effect of Treatment 1 is significant, indicating the existence of a control motive in the matching probabilities method. In Model C, we add two interaction terms, one checks whether the treatment had a differential effect depending on the Lab. We find no differences across Labs. The second term is an interaction between the percent probability of the lottery ticket M and treatment. Variable M continues to have a negative coefficient, but becomes significant, while the interaction with Treatment 2 is positive and significant too. This last effect is consistent with the alternative model we presented in Section 3.2, where subjects consider a lottery ticket itself to be the prize, rather than money. We do not have an explanation for why M is significant for T1, but it could be capturing some unobserved heterogeneity in subjects (say, some ability which makes subjects good at the visual task and in the quiz).

The table also presents, in the last column on the right, the analysis for all subjects. Even after including individuals for whom the mechanism was not incentive compatible (they had a success rate $s_i < 30\%$), the effect remains large and significant. As expected, the effect of Treatment 1 is smaller because subjects in Treatment 2 with $s_i < 30\%$ have a larger incentive to bet on themselves, since the probability of ticket M is capped at 100%.

The model predicts not only that $\bar{p}_1 > \bar{p}_2$, but also that the distribution of reported beliefs in Treatment 1 first order stochastically dominates the distribution of beliefs in Treatment 2. We explore this hypothesis in Figure 1, where we plot the cumulative distribution of placement by treatment. The cumulative distribution of p_1 lies below the one for p_2. This

[17] Subjects are not told their scores on the sample questions, but they probably formed beliefs about their performance in the sample. They are told the median score in previous sessions, and that the sample and quiz questions are similar. This enables them to transform their absolute inference into a relative belief.

Table 1: The effect of control on placement

	Pre-registered			Robustness
	Model A	Model B	Model C	All data
Treatment 1	9.672***	9.965***	20.53***	7.644***
	(0.000)	(0.000)	(0.000)	(0.000)
Munich		-1.228	-3.270	-0.668
		(0.552)	(0.255)	(0.732)
Sample Score		5.278***	5.332***	5.642***
		(0.000)	(0.000)	(0.000)
% probability ticket M		-0.0477	-0.175**	-0.0316
		(0.380)	(0.019)	(0.422)
Gender: Male		0.668	0.887	1.036
		(0.742)	(0.661)	(0.590)
Gender: Other		1.927	6.953	-2.352
		(0.909)	(0.681)	(0.842)
Economics		-1.868	-1.770	1.693
		(0.598)	(0.616)	(0.607)
PPLE		-3.818	-4.138	-0.348
		(0.496)	(0.460)	(0.946)
Other Social Sciences		-4.172	-4.183	-1.684
		(0.292)	(0.288)	(0.645)
STEM		0.690	0.627	3.368
		(0.921)	(0.928)	(0.597)
Treat.1 X Munich			4.058	
			(0.307)	
Treat.2 X % prob.M			0.261**	
			(0.013)	
Constant	54.54***	49.59***	43.89***	46.00***
N	536	536	536	610
R^2	0.0418	0.0799	0.0919	0.0699

Dependent variable: placement (report that will place in top half). Models A-C only include subjects for whom the elicitation is incentive compatible.

P-values in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

indicates that the result $\bar{p}_1 > \bar{p}_2$ is not due to only a handful of participants. Because the sample size needed to test for first order stochastic dominance with a reasonable degree of power was too large, we pre-registered that we expected the result of first order stochastic dominance, but that we did not expect to have enough power to reject equality of distributions. Nonetheless, the Mann-Whitney, or Wilcoxon Rank-Sum, test shows that we can reject equality of distributions: with a sample size of 271 for Treatment 1 and of 265 for Treatment 2, the z statistic is -4.17, which has a $p-$value of less than 1%. Accordingly, Kolmogorov-Smirnov and the median tests also have a $p-$value of less than 1%.

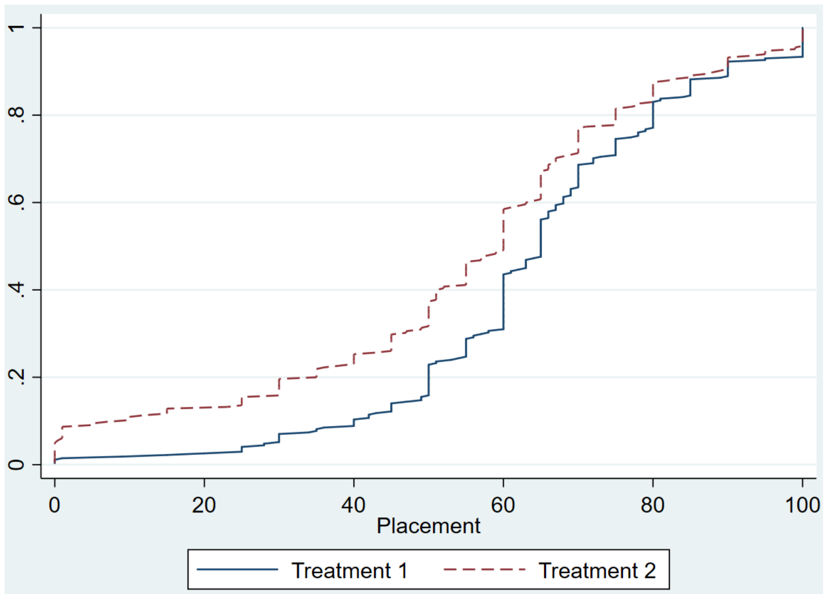


Figure 1. Cumulative distributions of p_1 and p_2.

5.1.1 A Replication of Owens et al.

Within the sample with scores in the visual task $s_i \geq \frac{3}{10}$, there were 15 subjects who did not complete an elicitation in at least one question; we keep them in the sample, but results are unchanged if we remove them. The average success rate in answering the questions was 42.9%, while the average reported chances of answering correctly were 51.5% in Treatment 2 and 54.2% in Treatment 1.

The main result of this section is that control inflates the average reported belief that a subject will answer correctly each of five questions. For each subject i we average her five reported beliefs to obtain q_i. The empirical strategy is as with p_i in the main experiment: regress q_i only on Treatment 1, and then on T1 and controls. The $t-$test of the regression of q_i on T1 gives a marginally significant value of 2.71 (one-sided $p-$value equal to 5.3%), but the regression with controls shows that in Treatment 1 subjects inflate their average responses by 3 percentage points, and that this coefficient is significant (one sided $p-$value $= 3\%$).

Our theory also predicts that the data cdf of q_i for Treatment 1 will first order stochastically dominate that of Treatment 2. Figure 2 shows that although the result does not obtain fully, it holds for average reports below 75%. We had pre-registered that we did not expect enough power to reject equality of distributions. In this case, unlike in the main analysis of Experiment 1, the results are only marginally significant. The $p-$value of the Wilcoxon rank-sum test is 9%, and that of the Kolmogorov-Smirnov test is 14%. So in this case the evidence in favor of our theory is weaker.

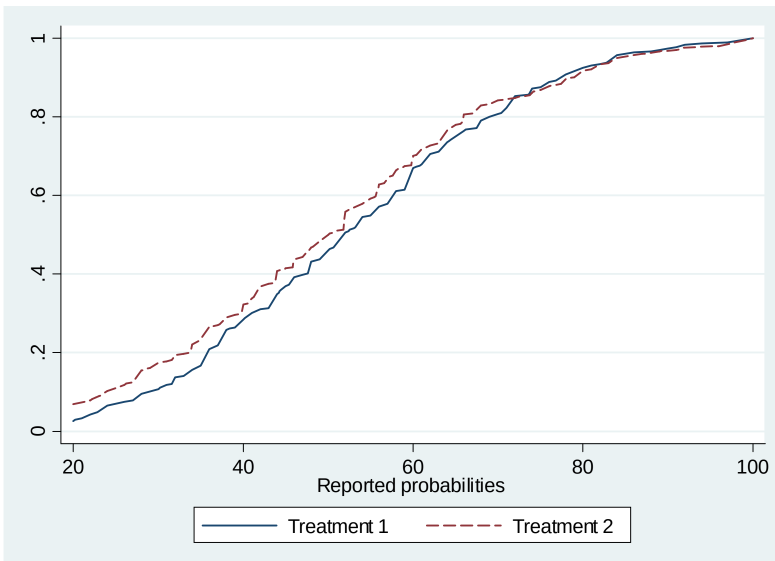


Figure 2. Cumulative distributions of q_1 and q_2.

5.2 Experiment 2

Four subjects did not report monotonic choices in the choice-lists and are excluded from the analysis.

The three treatments exhibit basically the same average estimate of p_i. In Treatment 1, with 66 subjects, $\bar{p}_1 = 66.61\%$; in Treatment 2, with 61 subjects, $\bar{p}_2 = 67.96\%$; in Treatment 3, with 65 subjects, $\bar{p}_3 = 66.45\%$. There are large standard deviations of comparable magnitude across treatments (16.90, 17.73 and 19.91 for Treatments $1-3$ respectively).

We perform two tests. With the Mann-Whitney test, the *p-value* for equality of distributions is 0.70 for Treatments 1 and 2, 0.70 for treatments 2 and 3, and 0.94 for Treatments 1 and 3. We also run the corresponding *ttest* for difference of means and we do not reject equality (*p-value* $= 0.66$ for Treatments 1 and 2, 0.66 for Treatments 2 and 3, and 0.96 for

Table 2: Effect of control on reported chance of answering questions correctly.

	Pre-registered			Robustness
	Model A	Model B	Model C	All data
Treatment 1	2.675	3.019*	10.76**	2.455*
	(0.105)	(0.060)	(0.017)	(0.098)
Munich		-2.973*	-6.854***	-3.085**
		(0.074)	(0.003)	(0.047)
Sample Score		4.968***	5.026***	5.141***
		(0.000)	(0.000)	(0.000)
% probability ticket M		-0.00289	-0.119**	-0.0348
		(0.947)	(0.046)	(0.266)
Gender: Male		3.160*	3.399**	3.465**
		(0.053)	(0.036)	(0.023)
Gender: Other		16.56	21.10	4.622
		(0.222)	(0.119)	(0.621)
Economics		5.596**	5.877**	6.076**
		(0.050)	(0.038)	(0.020)
PPLE		4.711	4.651	6.370
		(0.297)	(0.299)	(0.119)
Other Social Sciences		1.497	1.578	2.669
		(0.638)	(0.617)	(0.357)
STEM		10.86*	11.14**	11.34**
		(0.053)	(0.046)	(0.025)
Treat.1 X Munich			7.704**	
			(0.016)	
Treat.2 X % prob.M			0.238***	
			(0.005)	
Constant	51.40***	37.96***	33.50***	38.55***
N	536	536	536	610
R^2	0.00492	0.0821	0.105	0.0928

Dependent variable: average chance (reported average belief that previewed questions would be answered correctly). Models A-C only include subjects for whom the elicitation was incentive compatible.

P-values in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Treatments 1 and 3).

6 Conclusion

Social scientists are interested in people's beliefs. One way to elicit these beliefs is simply to ask for them. However, with little at stake people may provide ready answers with little connection to their actual beliefs. To counter this possibility, researchers have designed incentive schemes where people maximize their utility of money by reporting their beliefs. Yet these schemes remain vulnerable to distortions when subjects care about more than money. In particular, a desire for control may lead subjects to skew their reported beliefs under some ostensibly incentive compatible mechanisms. In an experiment using the matching probabilities method, we find that subjects inflate their reported beliefs about themselves by 18% for control benefits; non-monetary considerations account for at least 68% of what would otherwise be estimated to be overconfidence.

Although overconfidence and a desire for control can look similar to an observer, they are distinct phenomena with different implications. An overconfident person might quit her current job because she overestimates her prospects in another job with the same level of control; a well-calibrated CEO may make an acquisition simply because it yields her more control. Overconfidence is an error in beliefs; control is a preference. While supplying people with better information about themselves might change the behaviour of overconfident individuals, such information is irrelevant for choices that reflect a desire for control. A large body of research explores heterogeneity in overconfidence. People tend to be well-calibrated in their areas of expertise (see Schattka and Muller (2008)) and manifest little overplacement on attributes that are specific and objectively measured (Moore (2007)). No such links have been made to heterogeneity in control preferences.

Our study of control differs from earlier ones in that we introduce a new mechanism design that mitigates the control bias. Using this design, we still find measured overconfidence but the size is small: reported beliefs of placing in the top half average to 54% instead of 50%.

The new design can be used in contexts beyond the study of overconfidence.

7 Appendix A. Instructions for Experiment 1

Explanatory comments from the authors are, at times, interspersed among the instructions. They are indicated by use of the italic font and enclosed in square brackets.

Instructions

Welcome! This is an experiment in decision-making. If you follow the instructions and make good decisions you will earn a substantial amount of money. The money you earn will be paid to you at the end of the experiment. The experiment has four parts and there is a show-up fee of 7 euro that you will earn regardless of your choices. If you have questions throughout the experiment, you can send a message to the experimenter at any time via email (g.romagnoli@uva.nl).

Part 1

In the first part, you will perform 12 repetitions of the following exercise. First, you will see a string of numbers on the screen. Next, from memory you will type the numbers into the box appearing on the screen.

The time the numbers appear on the screen and the length of the string will vary across periods. Hence, remembering the string will be easier in some periods and harder in others. You will begin with two practice rounds and then repeat this exercise 10 times for payment.

Payment for this part:

At the end of the experiment, one round will be selected at random. If in that round you reported the string of numbers correctly, you will earn 2 euro for your correct answer.

Click on the Next button to proceed to the two sample rounds.

NOTE: The first blink starts immediately after you click on the next button.

Visual Task - Beginning of the 10 rounds used for payment

The two sample rounds are over. You will now play the next 10 rounds for payment.

Payment for this part (reminder):

At the end of the experiment, one round will be selected at random. If in that round you reported the string of numbers correctly, you earn 2 euro, otherwise you earn 0 euro.

End of part 1

The visual task is completed. You answered **N out of 10 rounds** correctly. *[Note: in the experimental screen, N is replaced by personalized values]*

Click on the Next button to proceed to the second part of the experiment.

Part 2

In this part, you are asked to answer a logic Quiz. The Quiz consists of 12 multiple-choice questions and you have 6 minutes overall to answer these questions.

Self-assessment:

Before you take the Quiz, we will ask you to estimate how likely you think it is that you will do better than half of the participants.

This is how we will rank the participants: After the quiz is completed, you will be assigned a position according to how many questions you answered correctly. The best performer among you will be assigned to rank 1, the second to rank 2, and so on. We will then list participants from the highest rank to the lowest rank and divide the subject pool into two equally sized-groups, a top half and a bottom half. For example, with 30 subjects the best 15 will be ranked in the top half and the other 15 will be ranked in the bottom half. Ties will be broken randomly. For instance, if two people are tied for 15th in terms of performance, then one of them will be randomly placed in the top half and one of them in the bottom half.

We want you to tell us your best estimate of the probability that you will place in the top half of the scores distribution. Your answer to this question will be measured in **chances**, which go from **0** (standing for: I am absolutely sure that my score will not be in the top half of the distribution) to **100** (standing for: I am absolutely sure that my score will be in the top half of the distribution). So, for example, 50 means that there are exactly equal chances that you score in the top or the bottom half; 75 means that you have the same chances to be in the top half as are the chances that a white ball is drawn from a bag with 75 white balls and 25 blue balls, and so on.

Please review your understanding of “chances" by answering the questions in the box below.

1. What are the chances that a fair coin is flipped and it turns up Tails?

25; 50*; 75; 100

2. What are the chances that a white ball is drawn from a bag with 30 white balls and 70 blue balls?

15; 30*; 60; 70

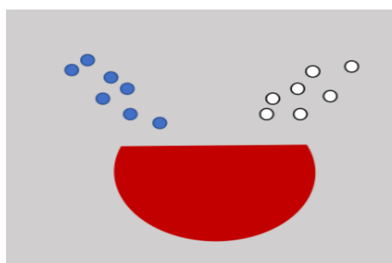
[The following portion of the instructions is different in the two treatments. Instructions for the two treatments are reported one after the other]

Treatment 1: Payment based on lottery tickets and BDM

Payment for this part:

We will ask you to report your chances of placing in the top half on the quiz. We follow a special procedure to reward you for reporting your chances as accurately as you can. This procedure is a bit complicated but the important thing to remember is that it is designed so that you will maximize your chances of winning a 20 euro payment by reporting your most accurate estimate of the probability that you place in the top half. The procedure is as follows.

On the screen, you can visualize a virtual bag. The bag is currently empty and will be filled with 100 blue and white balls. The exact composition of the virtual bag will be determined at the end of the experiment by a random device that will pick one among all the possible combinations [e.g., (0 white, 100 blue), (1 white, 99 blue), ... (49 white, 50 blue), (50 white, 50 blue), (51 white, 49 blue) , ... , (99 white, 1 blue), (100 white, 0 blue)], with equal likelihood.



There is a prize of a lottery ticket that gives a **30%** chance of winning 20 euro. You will have the opportunity to win the lottery ticket either (i) by betting that you place in the top half on the quiz, or (ii) by betting that a white ball is drawn from the virtual bag. If your chances of placing in the top half are greater than the number of white balls, then you will be more likely to win the 20 euro by betting on your placement, while if your chances of placing in the top half are less than the number of white balls, you will be more likely to win the 20 euro by betting on your performance on the virtual bag.

The exact procedure we use to determine your payment is as follows. First, we ask you to report what you think the chances **p** are that you place in the top half. Later, the computer will randomly pick a number of white balls for the virtual bag. Then,

- If your report **p** is greater than or equal to the number of white balls in the virtual bag, then **you will be betting that you place in the top half** : If your ranking on the quiz is in the top half, you will receive a lottery ticket that gives a **30%** chance of winning 20 euro. If your placement is in the bottom half, you will receive nothing.

- If instead the number of white balls in the virtual bag turns out to be greater than **p**, then **you will be betting on the virtual bag** : If a white ball is drawn from the virtual bag, you will receive a lottery ticket that gives a **30%** chance of winning 20 euro. If a blue ball is drawn, you will receive nothing.

Summing up, you might win the 20 euro either by placing in the top half or by a draw of a white ball from the virtual bag. With this procedure, you maximize the probability of winning the 20 euro by reporting your chances of placing in the top half as accurately as you can.

Please review your understanding of the payment procedure by answering the questions in the box below. NOTE: You can continue to the next page only after you answer correctly all the 5 questions on this page. You can try multiple times and there is no payment nor penalty for these review questions. If you need assistance, the experimenter is ready to help. You can contact her via email (g.romagnoli@uva.nl).

3. How is the composition of the virtual bag determined?

- The composition is determined randomly with some combinations of white and blue balls being more likely than others.
- The composition is predetermined by the experimenter.
- *The composition is determined randomly with each possible combination of white and blue balls having the same chance to be drawn.
- The composition is determined partly by chance and partly by my choices.

4. Anna reports a chance p=67% of placing in the top half and the computer draws a composition of 55 white balls and 45 blue balls for the virtual bag. Then:

- She always gets the lottery ticket irrespective of her choices.
- *Since her reported chances are larger than the number of white balls, she bets on her placement. She receives the lottery ticket if she placed in the top half of the score distribution.

- Since her reported chances are larger than the number of white balls, she bets on the virtual bag. She receives the lottery ticket if a white ball is drawn from the virtual bag.

- Since her reported chances are larger than the number of white balls, she obtains the lottery ticket.

5. Lisa reports a chance $p = 34\%$ of placing in the top half and the computer draws a composition of 41 white balls and 59 blue balls for the virtual bag. Then:

- Since her reported chances are smaller than the number of white balls, she bets on her placement. She receives the lottery ticket if she placed in the top half of the score distribution.

- She always gets the lottery ticket irrespective of her choices.

- *Since her reported chances are smaller than the number of white balls, she bets on the virtual bag. She receives the lottery ticket if a white ball is drawn from the virtual bag.

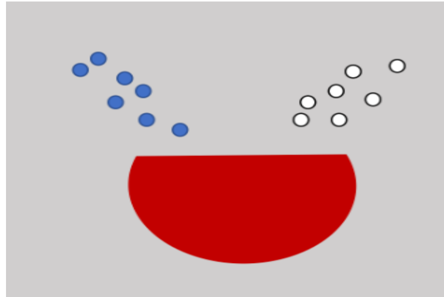
- Since her reported chances are smaller than the number of white balls, she obtains the lottery ticket.

Treatment 2: Payment based on lottery tickets and VisualTask-BDM

Payment for this part:

We will ask you to report your chances of placing in the top half on the quiz. We follow a special procedure to reward you for reporting your chances as accurately as you can. This procedure is a bit complicated but the important thing to remember is that it is designed so that you will maximize your chances of winning a 20 euro payment by reporting your most accurate estimate of the probability that you place in the top half. The procedure is as follows.

On the screen, you can visualize a virtual bag. The bag is currently empty and will be filled with 100 blue and white balls. The exact composition of the virtual bag will be determined at the end of the experiment by a random device that will pick one among all the possible combinations [e.g., (0 white, 100 blue), (1 white, 99 blue), ... (49 white, 50 blue), (50 white, 50 blue), (51 white, 49 blue) , ... , (99 white, 1 blue), (100 white, 0 blue)], with equal likelihood.



You will have the opportunity to win one of two lottery tickets for a prize of 20 euro. You will either (i) try to win Lottery Ticket A by betting that you place in the top half on the quiz, or (ii) try to win Lottery Ticket B by betting that you were successful on a random draw of one of your ten attempts of the visual task. Lottery Ticket A will give you 20 euro with a 30% chance if you place in the top half on the quiz; Lottery Ticket B will give you 20 euro with a $M\%$ chance, if you were successful in a random round of the visual task and a white ball is drawn from the virtual bag. *[in the experimental screen, M is replaced by personalized values* $min\{\frac{300}{10s_i}, 100\}$*]*. The lottery tickets have been specially calibrated so that: If your chances of placing in the top half are greater than the number of white balls, then you will be more likely to win the 20 euro by betting on your placement, while if your chances of placing in the top half are less than the number of white balls, then you will be more likely to win the 20 euro by betting on your performance on the visual task.

The exact procedure we use to determine your payment is as follows. First, we ask you to report what you think the chances **p** are that you place in the top half. Later, the computer will randomly pick a number of white balls for the virtual bag. Then,

- If your report **p** is greater than or equal to the number of white balls in the virtual bag, then **you will be betting that you place in the top half**. Specifically, if your placement on the quiz is in the top half, you will receive Lottery ticket A, which gives a **30%** chance of winning 20 euro. If your placement is in the bottom half, you will receive nothing.

- If your reported **p** is less than the number of white balls in the virtual bag, then **you will be betting on your skill on the visual task**: One of the 10 rounds that you completed in the visual task will be picked at random (with each round having the exact same probability of being selected) and a ball will be randomly drawn from the virtual bag. If you were successful at the visual task in the extracted round and the drawn ball is white, then you will receive Lottery ticket B, that gives you a **M%** chance of winning 20 euro. Otherwise you will receive nothing. *[in the experimental screen, M is replaced by personalized values* $min\{\frac{300}{10s_i}, 100\}$*]*

Summing up, you might win the 20 euro either by placing in the top half or by a successful performance on the visual task. With this procedure, you maximize the probability of winning the 20 euro by reporting your chances of placing in the top half as accurately as you can.

Please review your understanding of the payment procedure by answering the questions in the box below. NOTE: You can continue to the next page only after you answer correctly all the 5 questions on this page. You can try multiple times and there is no payment nor penalty for these review questions. If you need assistance, the experimenter is ready to help. You can contact her via email (g.romagnoli@uva.nl).

3. How is the composition of the virtual bag determined?

- The composition is determined randomly with some combinations of white and blue balls being more likely than others.
- The composition is predetermined by the experimenter.
- *The composition is determined randomly with each possible combination of white and blue balls having the same chance to be drawn.
- The composition is determined partly by chance and partly by my choices.

4. Anna reports a chance p=67% of placing in the top half and the computer draws a composition of 55 white balls and 45 blue balls for the virtual bag.

Then:

- *Since her reported chances are larger than the number of white balls, she bets on her placement. She receives a lottery ticket worth a 30% chance of winning 20 euro if she placed in the top half of the score distribution.
- She always gets a lottery ticket irrespective of her choices.
- Since her reported chances are larger than the number of white balls, she bets on her skill in the visual task. She receives ticket B if a white ball is drawn and she completed successfully a randomly selected round of the visual task.

- Since her reported chances are larger than the number of white balls, she obtains a lottery ticket worth a 30% chance of winning 20 euro.

5. Lisa reports a chance p=34% of placing in the top half, tickets A and B pay the prize with 30% and M% chance respectively, and the computer draws a composition of 41 white balls and 59 blue balls for the virtual bag. *[in the experimental screen, M is replaced by personalized values* $min\{\frac{300}{10s_i}, 100\}$*]*

   Then:

   - Since her reported chances are smaller than the number of white balls, she bets on her placement. She receives a ticket worth a 30% chance of winning 20 euro if she placed in the top half of the score distribution.

   - Since her reported chances are smaller than the number of white balls, she bets on the visual task. She receives ticket B if a white ball is drawn and she completed successfully a randomly selected round of the visual task.

   - Since her reported chances are smaller than the number of white balls, she obtains a lottery ticket worth a 30% chance of winning 20 euro.

   - She always gets the lottery ticket irrespective of her choices.

[The following instructions are common to both treatments, except when specified,]

Sample questions

Before you state your chances of placing in the top half, you will answer 3 sample questions which are comparable in difficulty to the questions that you will find in the Quiz. There is no payment for the sample questions but they give you an indication of the difficulty of the Quiz. You have 3 minutes in total to preview these questions. After they elapse, the page will auto-submit.

You are now ready to start the sample questions. Please click on the Next button.

Your assessment

What are your chances to be in the top half of the scores' distribution?

[Box for treatment 1]

At the end of the experiment, the computer will randomly pick a number of white balls for the virtual bag. Then,

- If your report **p** is greater than or equal to the number of white balls that end up in the virtual bag, then **you will be betting that you place in the top half**: If your ranking on the quiz is in the top half, you will receive a lottery ticket that gives a 30% chance of winning 20 euro. If your placement is in the bottom half, you will receive nothing.
- If instead the number of white balls in the virtual bag turns out to be greater than **p**, then **you will be betting on the virtual bag**: If a white ball is drawn from the virtual bag, you will receive a lottery ticket that gives a 30% chance of winning 20 euro. If a blue ball is drawn, you will receive nothing.

Before you report your chances, please review **the payment procedure for this part** in the box below.

[The boxes for Treatment 1 and Treatment 2 are reported one after the other]

Some further notes:

- The sample questions you just saw are of comparable difficulty to the actual questions you will encounter in the Quiz.
- In past sessions, the better performing half of the subjects answered **7** or more questions correctly, out of a total of 12 questions.

You are now ready to state your chances to be in the top half of the scores distribution.
Type a number between **0** (meaning: I have zero chance to be in the top half) to **100** (meaning: I am absolutely sure I will be in the top half of score distribution).

My chances:

NOTE: The Quiz starts immediately after you click on the next button. You have 6 minutes to answer the 12 questions. After the time elapses, the page will submit automatically.

[Box for treatment 2]

At the end of the experiment, the computer will randomly pick a number of white balls for the virtual bag. Then,

- If your report **p** is greater than or equal to the number of white balls in the virtual bag, then **you will be betting that you place in the top half**. Specifically, if your placement on the quiz is in the top half, you will receive Lottery ticket A, which gives a 30% chance of winning 20 euro. If your placement is in the bottom half, you will receive nothing.

- If your reported **p** is less than the number of white balls in the virtual bag, then **you will be betting on your skill on the visual task**: One of the 10 rounds that you completed in the visual task will be picked at random (with each round having the exact same probability of being selected) and a ball will be randomly drawn from the virtual bag. If you were successful at the visual task in the extracted round and the drawn ball is white, then you will receive Lottery ticket B, that gives you a $M\%$ chance of winning 20 euro. Otherwise you will receive nothing. *[in the experimental screen, M is replaced by personalized values* $min\{\frac{300}{10s_i}, 100\}$*]*

End of the quiz - Transition to Part 3

The Quiz is completed and Part 2 is over.
Please, click on the next button when you are ready to move to the third part of the experiment.

Part 3

This third part is in preparation for Part 4, in which you will be presented with 5 multiple-choice logic questions which you will solve for a financial reward. For each of these questions, you will have 45 seconds to provide the answer. In this third part, you are asked to predict the chances that you will answer each of these questions correctly.
You make this prediction after previewing each of the questions for 15 seconds. You will start with the preview of question 1. After the 15 seconds elapse, the question will disappear and you will be asked to estimate how likely you will be to answer this question correctly once it is presented again in Part 4.
This time, the minimum chance you can report is **20**. This is because, when you have

absolutely no idea of what the right answer is, you will guess one option at random. Since each question presents you with 5 options, a random guess will have 1 in 5 or, equivalently, a 20% chance to be correct. Thus, the range of chances you can report goes from **20** (meaning: I will have a completely random, 20% chance, to be correct) to **100** (meaning: I am absolutely sure I will answer correctly).

You will have a maximum of 25 seconds to enter your chances. After you have entered your chances of a correct answer, please press the "Next" button or wait for the 25 seconds to elapse in order to continue to the next question. The second through fifth questions will proceed in the same manner.

Payment for this part

For each assessment you give of your chances of answering a question correctly, we will follow a procedure identical to the one explained in Part 2 to give you the possibility of winning a lottery ticket (except for the prize, which is 10 euro). At the end of the experiment, one of the five assessments is selected at random and the following procedure is implemented:

A second virtual bag will be filled with a random composition of blue and white balls as explained in Part 2 (the composition of the second virtual bag is randomly drawn and independent from the composition of the first virtual bag used for Part 2).

[The following portion of the instructions is again different in the two treatments. Instructions for the two treatments are reported one after the other]

[Treatment 1: Payment based on lottery tickets and BDM]

You will have the opportunity to win the lottery ticket either (i) by betting that your answer in Part 4 is correct or (ii) by betting that a white ball is drawn from the virtual bag. Given your reported chances p that your answer to the question will be correct, your payment will be determined as follows:

- If **p** is greater or equal than the number of white balls, then **you will be betting on your answer being correct** . That is, if you answered this question correctly you will receive a lottery ticket that gives a 30% chance of winning 10 euro. If you answered incorrectly, you will receive nothing.

- If instead your reported **p** is strictly less than the number of white balls in the virtual bag, then **you will be betting on the virtual bag**. That is, a ball will be drawn from the virtual bag and if it is white you will receive a lottery ticket that gives a 30% chance of winning 10 euro. If a blue ball is drawn, you will receive nothing.

[Treatment 2: Payment based on lottery tickets and VisualTask-BDM]

You will have the opportunity to win one of two lottery tickets either (i) by betting on your answer being correct; or (ii) by betting that you were successful on a random draw of one of your ten attempts of the visual task.
Given your reported chances **p** that your answer in Part 4 will be correct, your payment will be determined as follows:

- If **p** is greater than or equal to the number of white balls in the virtual bag, then you will **bet that you answer the question correctly**: if your answer is correct, you will receive a lottery ticket that gives a 30% chance of winning 10 euro. If your answer is incorrect, you will receive nothing.

- If your reported **p** is strictly less than the number of white balls in the virtual bag, then **you will be betting on your skill on the visual task**: Specifically, one of the 10 rounds that you completed in the visual task will be picked at random (with each round having the exact same probability of being selected) and a ball will be randomly drawn from the virtual bag. If you were successful at the visual task in the extracted round and the drawn ball is white, then you will receive Lottery ticket B, which gives you a M% chance of winning 10 euro. Otherwise you will receive nothing.*[Note: in the experimental screen, M is replaced by personalized values* $min\{\frac{300}{10s_i}, 100\}$*]*.

[The following instructions are common to both treatments]

NOTE: All remaining pages before the end of the experiment are characterized by automatic and timed transitions. You will first preview each question for 15 seconds and make your assessment, then you will automatically transition to Part 4 where you answer each of the previewed questions within 45 seconds for each question. **Thus, before clicking on Next, please make sure that you have 8 uninterrupted minutes ahead of you.**

Question 1 *[to 5]*: Your chances

What are your chances that you will answer the question you just previewed correctly?

Type a number between **20** (meaning: I will have a completely random, 20% chance, to be correct) to **100** (meaning: I am absolutely sure I will answer correctly).

My chances:

Part 4

Part 3 is complete. You are now moving to Part 4.
You will see again the 5 questions you have just previewed, one by one. Your task is to find the correct answer among the 5 presented for each question. You will have 45 seconds per question.

Payment for this part

One among the five questions will be selected at random, with each question having equal chances to be drawn. If your answer in that question was correct you will receive 2 euro.

End of the experiment - Demographics

The fourth and last part of the experiment has concluded.
Please fill up the demographic questions below. Afterwards, you can click on the Next button and review your total payoff.

8 Appendix B. Instructions for Experiment 2

We present instructions for Experiment 2. Explanatory comments from the authors are, at times, interspersed among the instructions. They are indicated by use of the italic font and enclosed in square brackets.

Instructions

This is an experiment in decision making. Funds have been provided to run this experiment. If you follow the instructions and make good decisions, you will earn a substantial amount of money. The money you earn will be paid to you in CASH at the end of the experiment. The session will take place through computer terminals. There is a show-up fee of 10 euro that you will earn regardless of your choices. The experiment will consist of two parts. At the end of the experiment, a random device will determine whether you are going to be paid according to your answers in the first part or in the second part of the experiment, with a 50% chance that each part is used for payment.

Please turn off your phones now and do not talk or communicate to each other in any way.

First part

In the first part of the experiment, you are asked to answer a logic quiz. The quiz consists of 20 multiple-choice questions and you have 13 minutes to answer the questions. You will earn 50 cents for each correct answer and zero cents for each incorrect answer. Hence, if this first part of the experiment is randomly drawn and used for payment, you can earn from a minimum of 10 euro to a maximum of 20 euro including the show-up fee.

[The second part is presented separately for each of the 3 treatments].

Second part (Treatment 1 - Betting up)

In this second part of the experiment, we ask you to estimate how well you did in the quiz relative to the other subjects. Of course, you cannot know your relative performance for sure so we will ask you for a probability estimate. Specifically, we will ask you with which probability you think you placed in the upper half of subjects.

You will be assigned a ranking based on how many questions you answered correctly in the quiz you just took. The best performer among you will be assigned to rank 1, the second best performer to rank 2 and so on. We will then list the participants in the experiment from the highest rank to the lowest rank and divide the subject pool into two equally sized-groups, an upper half and a lower half. For example, with 14 subjects the top 7 will be ranked in the upper half and the other seven will be ranked in the lower half. If, say, two people are tied for 7^{th} in terms of performance, then one of them will be randomly placed in the upper half and one of them in the lower half.

We want you to tell us your best estimate of the probability that you are in the upper half. For this purpose, we will use a special payment procedure that rewards you for giving us your best estimate. The procedure is a bit complicated but the most important thing to understand about it is simply that you maximize your expected payment by reporting your best estimate. We now explain this procedure.

At the end of the experiment, the computer will create a virtual bag. The bag will be filled with 100 blue and white balls. The exact composition of the virtual bag will be determined at the end of the experiment by a random device that will pick one of the following possibilities with equal likelihood: (0 white, 100 blue), (2 white, 98 blue), (4 white, 96 blue) ... (98 white, 2 blue), (100 white, 0 blue) - so the virtual bag will have one among all possible combinations of white and blue balls with increments of two.

There is a prize of 10 euro that you have a chance to win by either betting on your placement or by betting on the virtual bag. For each of the possible combinations, we want to know if you prefer to bet on your placement or to bet on a white draw from the virtual bag. Choices will be presented to you in a list of pairwise comparisons, as shown in Figure 1.

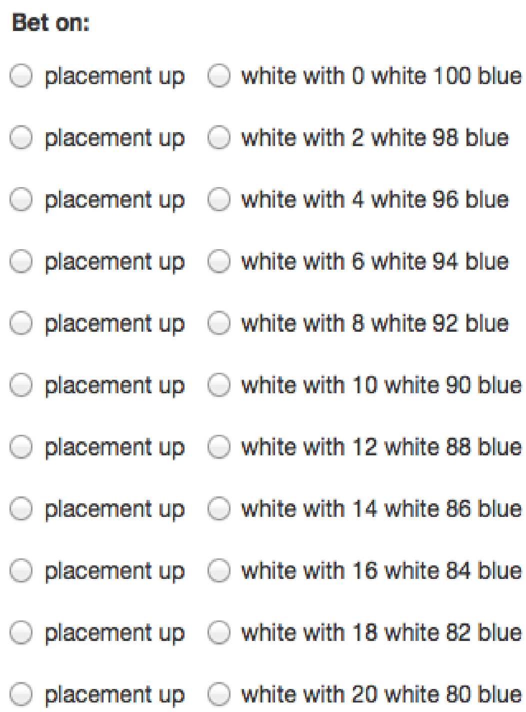


Figure 1. Choices

In each comparison you choose between betting on your placement-up or on the virtual bag:

- If you bet on your **placement-up,** you win 10 euro if you are in the upper half of the ranking and 0 euro otherwise.
- If you bet on the **virtual bag**, you win 10 euro if a white ball is drawn from the virtual bag and 0 euro otherwise.

Thresholds: The number of white balls represents your chances of winning when you bet on the virtual bag. The number of white balls increases as you scroll down the list, so the virtual bag becomes more attractive the more down you go on the list. Hence we expect that, if you choose the virtual bag in one comparison, you will choose the virtual bag in all comparisons that follow below it. In other words, we expect that you will have a **threshold**, that is, a certain amount of white balls such that you bet on your placement-up until that threshold and then switch to bet on the virtual bag if it contains more white balls than the threshold. We will interpret this threshold as the probability that you believe your score falls in the upper half of the distribution.

You can try out different thresholds and your choice will be final only when you click on the Next button. Remember, once again, that you maximize your chances of winning if your threshold is the probability that you assign to having a quiz score in the upper half of the distribution.
At the end of the experiment, a random device will select one of the questions, that is, one of the possible bag compositions. Then one ball will be extracted from the virtual bag. Your payment will depend on the color of the ball and your choice in the selected question. To recap, if, in the selected question:

- You bet on the virtual bag, then you win 10 euro if a white ball is randomly extracted from the bag;
- You chose to bet on you placement-up, then you win 10 euro if you placed in the upper half.

Examples: Lisa thinks there is a 60% chance she placed in the upper half. Hence, she chooses to bet on her placement-up if in the bag there are 60 white balls or fewer and on the virtual bag if it contains more than 60 white balls. She, therefore, clicks all the buttons according to this rule and her choices will look as in Figure 2:

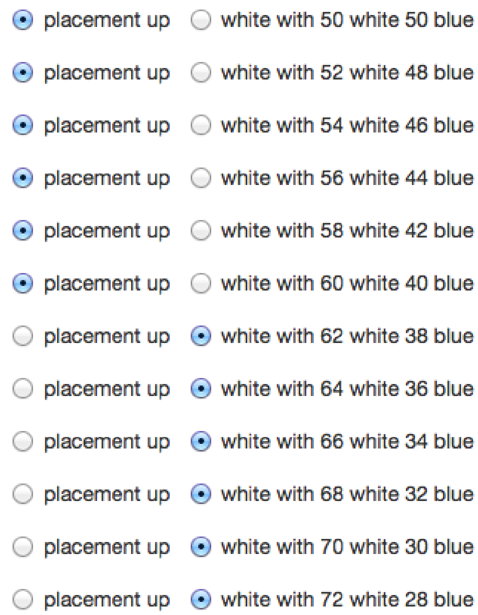


Figure 2. Lisa's Choices

John thinks there is a 20% chance he placed in the upper half. Hence, he chooses to bet on his placement-up if there are 20 white balls or fewer in the virtual bag, otherwise he prefers to bet on the virtual bag. He clicks the buttons according to this threshold and his choices will look as in Figure 3.

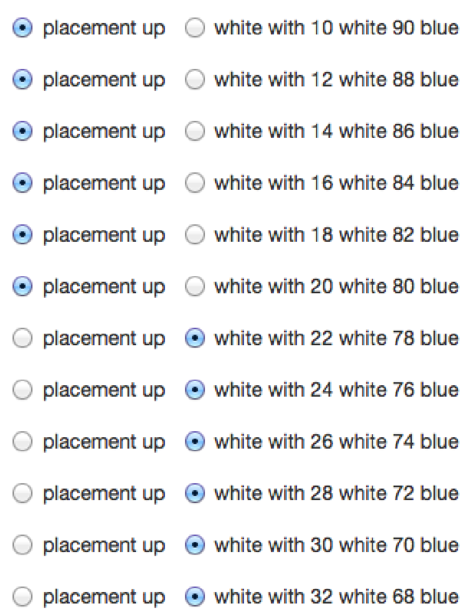


Figure 3. John's Choices

Second part (Treatment 2 - Betting down)

In this second part of the experiment, we ask you to estimate how well you did in the quiz relative to the other subjects. Of course, you cannot know your relative performance for sure so we will ask you for a probability estimate. Specifically, we will ask you with which probability you think you placed in the lower half of subjects.

You will be assigned a ranking based on how many questions you answered correctly in the quiz you just took. The best performer among you will be assigned to rank 1, the second-best performer to rank 2 and so on. We will then list the participants in the experiment from the highest rank to the lowest rank and divide the subject pool into two equally sized-groups, an upper half and a lower half. For example, with 14 subjects, the top 7 will be ranked in the upper half and the other seven will be ranked in the lower half. If, say, two people are tied for 7^{th} in terms of performance, then one of them will be randomly placed in the upper half and one of them in the lower half.

We want you to tell us your best estimate of the probability that you are in the lower half. For this purpose, we will use a special payment procedure that rewards you for giving us your best estimate. The procedure is a bit complicated but the most important thing to understand about it is simply that you maximize your expected payment by reporting your best estimate. We now explain this procedure.

At the end of the experiment the computer will create a virtual bag. The bag will be filled with 100 blue and white balls. The exact composition of the virtual bag will be determined at the end of the experiment by a random device that will pick one of the following possibilities with equal likelihood: (0 white, 100 blue), (2 white, 98 blue), (4 white, 96 blue) ... (98 white, 2 blue), (100 white, 0 blue) - so the virtual bag will have one among all possible combinations of white and blue balls with increments of two.

There is a prize of 10 euro that you have a chance to win by either betting on your placement or by betting on the virtual bag. For each of the possible combinations, we want to know if you prefer to bet on your placement or to bet on a white draw from the virtual bag. Choices will be presented to you in a list of pairwise comparisons, as shown in Figure 1.

Figure 1. Choices

In each comparison you choose between betting on your placement-down or on the virtual bag:

- If you bet on your **placement-down,** you win 10 euro if you are in the lower half of the ranking and 0 euro otherwise.

- If you bet on the **virtual bag**, you win 10 euro if a white ball is drawn from the virtual bag and 0 euro otherwise.

Thresholds: The number of white balls represents your chances of winning when you bet on the virtual bag. The number of white balls increases as you scroll down the list, so the virtual bag becomes more attractive the more down you go on the list. Hence we expect that, if you choose the virtual bag in one comparison, you will choose the virtual bag in all comparisons that follow below it. In other words, we expect that you will have a **threshold**, that is, a certain amount of white balls such that you bet on your placement-down until that threshold and then switch to bet on the virtual bag if it contains more white balls than the threshold. We will interpret this threshold as the probability that you believe your score falls in the lower half of the distribution.

You can try out different thresholds and your choice will be final only when you click on the Next button. Remember, once again, that you maximize your chances of winning if your

threshold is the probability that you assign to having a quiz score in the lower half of the distribution.
At the end of the experiment, a random device will select one of the questions, i.e. one of the possible bag compositions. Then one ball will be extracted from the virtual bag. Your payment will depend on the color of the ball and your choice in the selected question. To recap, if, in the selected question:

- You bet on the virtual bag, you win 10 euro if a white ball is randomly extracted from the bag;
- You chose to bet on you placement-down, you win 10 euro if you placed in the lower half.

Examples: Lisa thinks there is a 60% chance she placed in the lower half. Hence, she chooses to bet on her placement-down if in the bag there are 60 white balls or fewer and on the virtual bag if it contains more than 60 white balls. She, therefore, clicks all the buttons according to this rule and her choices will look as in Figure 2:

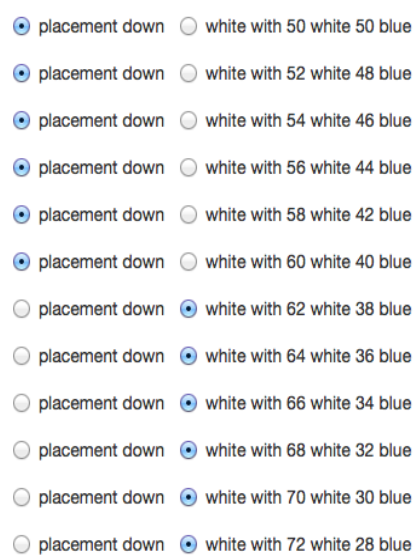


Figure 2. Lisa's Choices

John thinks there is a 20% chance he placed in the lower half. Hence, he chooses to bet on his placement-down if there are fewer than 20 white balls in the virtual bag, otherwise he prefers to bet on the virtual bag. He clicks the buttons according to this threshold and his choices will look as in Figure 3.

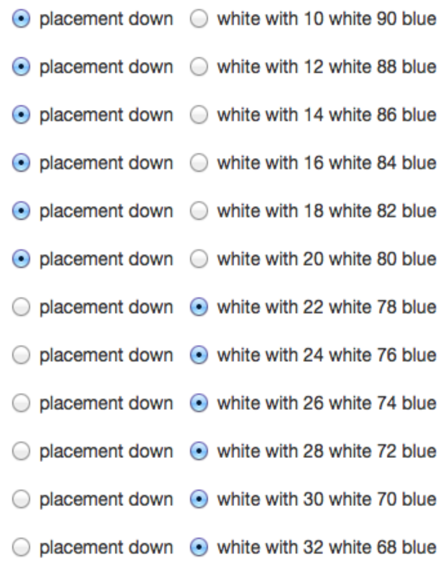


Figure 3. John's Choices

Second part (Treatment 3 - Betting up and down)

In this second part of the experiment, we ask you to estimate how well you did on the quiz relative to the other subjects. Of course, you cannot know your relative performance for sure so we will ask you for a probability estimate. Specifically, we will ask you with which probability you think you placed in the lower half of subjects.[18]

You will be assigned a ranking based on how many questions you answered correctly on the quiz you just took. The best performer among you will be assigned to rank 1, the second best performer to rank 2 and so on. If there are ties, these ties will be broken randomly, so that everyone is assigned a unique rank.

We will then list the participants in the experiment from the highest rank to the lowest rank and divide the subject pool into two equally sized-groups, an upper half and a lower half. For example, with 14 subjects the top 7 will be ranked in the upper half and the other seven will be ranked in the lower half. If, say, two people are tied for 7^{th} in terms of performance, then one of them will be randomly placed in the upper half and one of them in the lower half.

We want you to tell us your best estimate of the probability that you are in the lower half. For this purpose, we will use a special payment procedure that rewards you for giving

[18] *[Note: In this treatment, subjects bet on both their performance being in the upper part and in the lower part of the distribution. In 2 (out of 4) sessions, the framing of the instructions starts off with betting-down and later introduces betting-up, in the other two treatments the order in which the two types of bets are presented is reversed].*

us your best estimate. The procedure is a bit complicated but the most important thing to understand about it is simply that you maximize your expected payment by reporting your best estimate. We now explain this procedure.

At the end of the experiment, the computer will create a virtual bag. The bag will be filled with 100 blue and white balls. The exact composition of the virtual bag will be determined at the end of the experiment by a random device that will pick one of the following possibilities with equal likelihood: (0 white, 100 blue), (2 white, 98 blue), (4 white, 96 blue) ... (98 white, 2 blue), (100 white, 0 blue) - so the virtual bag will have one among all possible combinations of white and blue balls with increments of two.

There is a prize of 10 euro that you have a chance to win by either betting on your placement or by betting on the virtual bag. For each of the possible combinations, we want to know if you prefer to bet on your placement or to bet on a white draw from the virtual bag. Choices will be presented to you in two groups of pairwise comparisons, as shown in Figure 1.

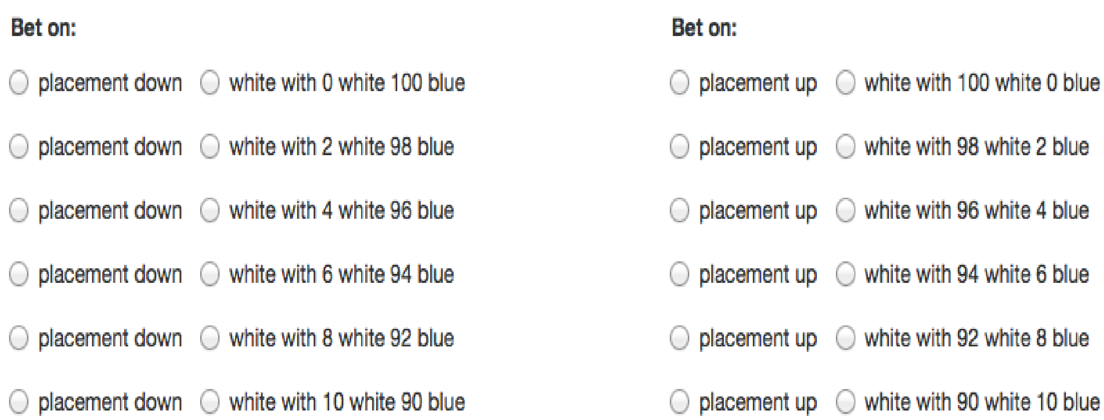


Figure 1. Choices

In the column on the left, you choose between betting on your placement-down or on the virtual bag:

- If you bet on your **placement-down,** you win 10 euro if you are in the lower half of the ranking and 0 euro otherwise.
- If you bet on the **virtual bag**, you win 10 euro if a white ball is drawn from the virtual bag and 0 euro otherwise.

In the column on the right, you choose between betting on your placement-up or on the virtual bag:

- If you bet on your **placement-up,** you win 10 euro if you are in the upper half of the ranking and 0 euro otherwise.
- If you bet on the **virtual bag,** you win 10 euro if a white ball is drawn from the virtual bag and 0 euro otherwise.

[Note: In 2 (out of 4) sessions of treatment 3, the order of the columns was reversed and the instructions were adjusted accordingly. As a result, subjects would bet on their placement-up in the column on the left, and on their placement-down in the column on the right.]

Thresholds: The number of white balls represents your chances of winning when you bet on the virtual bag. In the left column, the number of white balls increases as you scroll down the list, so the virtual bag becomes more attractive the more down you go on the list. Hence we expect that, if you choose the virtual bag in one comparison, you will choose the virtual bag in all comparisons that follow below it. In other words, we expect that you will have a **threshold**, that is, a certain amount of white balls such that you bet on your placement-down until that threshold and then switch to bet on the virtual bag if it contains more white balls than the threshold. We will interpret this threshold as the probability that you believe your score falls in the lower half of the distribution.

Your choices from the right column. In the right column, you are choosing between betting on your placement-up or the virtual bag. Here the order of the virtual bags is reversed: The number of white balls starts at 100 and decreases as you scroll down the list. Here again, you will have a threshold: You will start betting on the virtual bag and then switch at some point to betting on your placement-up. This threshold will tell us the probability with which you believe your score belongs to the upper half of the distribution.

Admissible choices: The choices from the two columns are tied together, that is, **the two thresholds will have to be placed on the same line**. The reason is that if you told us that there is an $x\%$ chance that your rank is in the lower half, we will presume you think there is a $100 - x\%$ chance that your score is in the upper half. In Figure 5, you can see a preview of what it means for the two thresholds to be placed on the same line. We'll go back to it at the end.

A way to ensure you are meeting this constraint is to verify that, taking two questions placed on the same line, you are betting on the virtual bag in one and only one of them. Figure 2 shows two examples of non-admissible choices. Figure 3 shows two examples of admissible choices. If you make a mistake, an error message will prompt you to correct your entries until only admissible choices are present.

Figure 2. Non-admissible choices

Figure 3. Admissible choices

You can try out different thresholds and your choice will be final only when you click on the Next button. Remember, once again, that you maximize your chances of winning if, in the left column, your threshold is the probability that you assign to having a quiz score in the lower half of the distribution, and, in the right column, you pick as threshold the probability that your score is in the upper half.
At the end of the experiment, a random device will select one of the two groups of questions and one of the possible bag compositions. Then one ball will be extracted from the virtual bag. Your payment will depend on the color of the ball and your choice in the selected question. To recap, if in the selected question:

- You bet on the virtual bag, you win 10 euro if a white ball is randomly extracted from the bag;
- You chose to bet on your placement-down, you win 10 euro if you placed in the lower half;
- You chose to bet on your placement-up, you win 10 euro if you placed in the upper half.

Examples: Lisa thinks there is a 60% chance she placed in the lower half and a 40% chance she placed in the upper half. Hence, she chooses to bet on her placement-down if in the bag there are 60 white balls or fewer. Moreover, she chooses to bet on her placement-up if there are fewer than 40 white balls in the bag. She therefore clicks all the buttons according to this rule and her choices will look as in Figure 4:

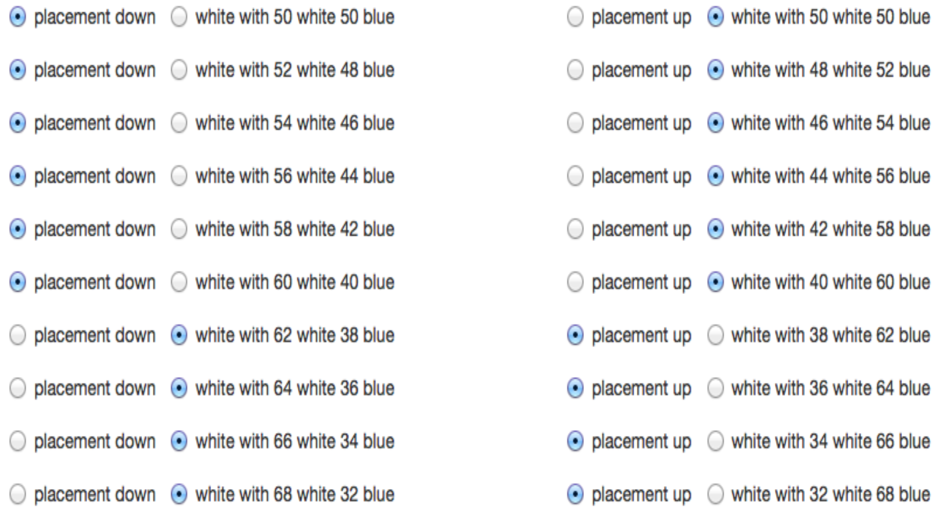


Figure 4. Lisa's Choices

John thinks there is a 20% chance he placed in the lower half. Hence, he chooses to bet on his placement-down rather than on the virtual bag if there are fewer than 20 white balls in the virtual bag, otherwise he prefers to bet on the bag. He clicks the buttons according to this threshold and his choices will look as in Figure 5. This should be consistent with his belief that there is an 80% probability that he scored in the upper half.

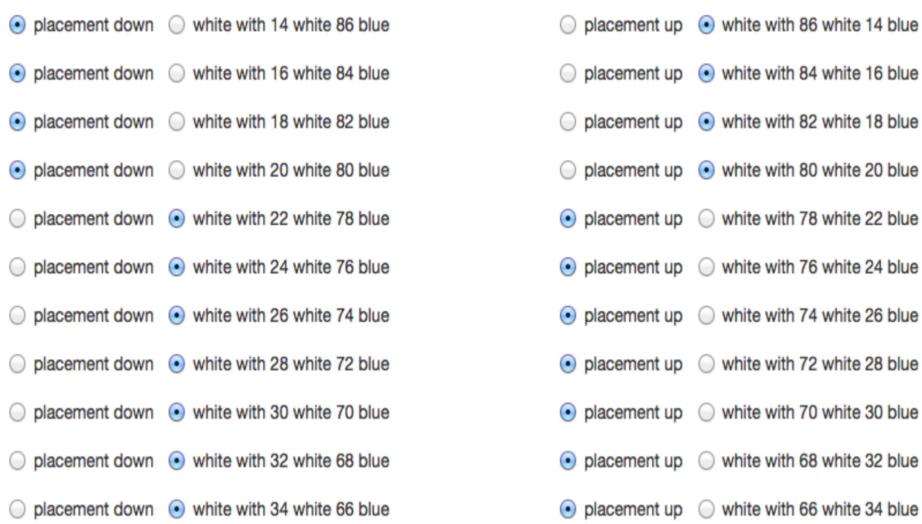


Figure 5. John's Choices

9 Appendix C. Pre-Registration.

AS PREDICTED

Effect of control on Elicitation 2020 (#41354)

Created: 05/19/2020 01:07 PM (PT)
Public: 05/19/2020 04:25 PM (PT)

Author(s)
Jean-Pierre Benoît (London Business School) - jpbenoit@london.edu
Juan Dubra (Universidad de Montevideo) - dubraj@um.edu.uy
Giorgia Romagnoli (University of Amsterdam) - g.romagnoli@uva.nl

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
We ask to what extent control (the desire to bet on activities where one could in principle affect the outcome) affects the self-reported estimated probabilities of success, when these are elicited with the matching probabilities rule of Grether (1981).

3) Describe the key dependent variable(s) specifying how they will be measured.
a) The dependent variable is the self-reported belief of the likelihood that the subject performs better than half of the other subjects in a quiz. We will compare two measures: in Treatment 1 we elicit beliefs with the matching probabilities rule (with payment in lottery tickets); in Treatment 2 beliefs are elicited with a modified version of the matching probabilities rule where the outcome is always dependent on the performance of the subject (either in the main task or in a secondary visual/memory task performed in the computer).
b) In the second part of the experiment, subjects are shown a multiple choice question for a few seconds. Then they are asked to predict the probability they think they have of answering the question correctly. This is repeated for five questions. The elicitation of the beliefs is done, as in part (a) with two different mechanisms: the Grether-cum-lottery ticket or the variation in which subjects bet on their performance (either on answering the question correctly or in the visual/memory task).

4) How many and which conditions will participants be assigned to?
We will have two treatments (with around 300 participants in each; they will be from the CREED Lab in the Netherlands and from Munich, with both labs running both treatments) corresponding to the two different ways of measuring beliefs as explained in section 3.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
a) We will perform two OLS regressions: one with belief as the dependent variable, and an indicator for Treatment 1 a dummy; in the other we will add controls such as Gender, probability of winning the lottery, sample score, major, and lab. We expect the coefficient in the dummy variable to be positive; we will report one sided tests.
b) We will first average the five elicited probabilities of answering each question correctly for each subject. Then we will perform the same analysis as in (a). Again, we expect a positive coefficient on the dummy "Treatment 1".

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
We will exclude people who score less than 3 in the visual task, and those who choose not to complete the experiment.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
Around 600 subjects will participate, depending on how experimental sessions are filled.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
We don't expect to have enough power to reject equality of distributions when plotting the cdf of responses by treatment, but we expect the distribution of Treatment 2 to be mostly below that of Treatment 1. Similarly, we don't expect to have power to reject that control is increasing in belief, but we will present two statistical analyses to illustrate: for (a) a regression of placement adding the interaction of sample score with treatment; for (b) control should be larger in easier questions, so control in the two questions perceived as easier (by average report in T2) should be larger than in the two harder ones.

References

[1] Allen, F, (1987), "Discovering Personal Probability when Utility Functions Are Unknown," *Management Science*, 33(4), pp. 542-544.

[2] Andreoni, J. and A. Sanchez (2014), "Do Beliefs Justify Actions or Do Actions Justify Beliefs? An Experiment on Stated Beliefs, Revealed Beliefs, and Social-Image Motivation" mimeo.

[3] Arechar, A., S. Gächter and L. Molleman, (2018) "Conducting interactive experiments online," *Experimental Economics* 21, 99–131.

[4] Armantier, O. and N. Treich (2013), "Eliciting Beliefs: Proper Scoring Rule, Incentives, Stakes and Hedging" European Economic Review, Volume 62, 17-40, 2013

[5] Baumeister, R.F. (1982), "A self-presentational view of social phenomena," *Psychological Bulletin* 91:3–26.

[6] Benoît, J-P. and J. Dubra (2011), "Apparent Overconfidence" *Econometrica* **79(5)**, 1591-1625.

[7] Benoît, J.-P., Dubra, J. and Moore, D. A. (2015), Does the Better-than-average Effect show that People are Overconfident?: Two Experiments. Journal of the European Economic Association, 13: 293-329.

[8] Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2019), "Beliefs about Gender," *American Economic Review* 109(3): 739-773.

[9] Burks, S., J. Carpenter, L. Goette and A. Rustichini (2013), "Overconfidence and Social Signalling," Review of Economic Studies, 80(3), 949-83.

[10] Buser, T. L. Gerhards and J. van der Weele (2018) "Responsiveness to feedback as a personal trait," *Journal of Risk and Uncertainty* 56(2), 165–192.

[11] Cacault and Grieder (2019), "How Group Identification Distorts Beliefs," *Journal of Economic Behavior and Organization*, 164.

[12] Camerer, C., and Lovallo, D. (1999), "Overconfidence and Excess Entry: An Experimental Approach," The American Economic Review 89.1 : 306â€“318.

[13] Chen, D. L., Schonger, M. and Wickens, C. (2016), "oTree - An open-source platform for laboratory, online, and field experiments," J. Behav. Exp. Finance 9, 88-97.

[14] Chew S. H., and J. Sagi, (2008), "Small worlds: Modeling attitudes toward sources of uncertainty," *Journal of Economic Theory*, 139(1), pp. 1-24.

[15] Clark, Jeremy and Lana Friesen (2009), "Overconfidence in Forecasts of Own Performance: An Experimental Study," Economic Journal, 119(1), 229-51.

[16] Costello, F. and Watts, P. (2014), "Surprisingly Rational: Probability Theory Plus Noise Explains Biases in Judgment," *Psychological Review*, 121, 463-480.

[17] Coffman, K. M. Collis and L. Kulkarni (2019), "Stereotypes and Belief Updating", HBS Working Paper #19-068.

[18] Coutts, A. (2019) "Good News and Bad News are Still News: Experimental Evidence on Belief Updating," *Experimental Economics*, 22(2), 369-395.

[19] Ducharme, W. and Donnell, M. (1973), "Intrasubject Comparison of Four Response Modes for Subjective Probability Assessment," *Organizational Behavior and Human Performance* 10, 108-117.

[20] Ewers, M. and F. Zimmermann (2015), "Image and Misreporting," Journal of the European Economic Association, 13(2), 363–80.

[21] Goodie, A. S. (2003) The effects of control on betting: paradoxical betting on items of high confidence with low value. J Exp Psychol Learn Mem Cogn: 29, 598-610.

[22] Goodie, Adam and Diana Young (2007), "The skill element in decision making under uncertainty: Control or competence?," Judgment and Decision Making, 2(3), pp. 189-203.

[23] Grether, D. M. (1981) "Financial Incentive Effects and Individual Decision Making," Social Sciences working paper 401, Cal. Tech.

[24] Grether, D. M. (1992) "Testing Bayes Rule and the Representativeness Heuristic: Some Experimental Evidence." *Journal of Economic Behavior and Organization* **17**, 31–57.

[25] Gul, R and W. Pesendorfer (2015) "Hurwicz Expected Utility and Multiple Sources," Journal of Economic Theory.

[26] Heath, Chip and Amos Tversky, (1991) "Preference and Belief: Ambiguity and Competence in Choice under Uncertainty," Journal of Risk and Uncertainty, 4, 5-28.

[27] Hoelzl, E. and A. Rustichini, (2005), "Overconfident: do you put your money on it?" the *Economic Journal*, **115**, pp. 305-18.

[28] Hollard, G., Massoni, S., & Vergnaud, J.-C. (2016), "In Search of Good Probability Assessors: An Experimental Comparison of Elicitation Rules for Confidence Judgments," Theory and Decision. 80(3): 363-387.

[29] Hossain, T. and R. Okui (2013), "The Binarized Scoring Rule," Review of Economic Studies 80, 984-1001.

[30] Karni, E. (2009), "A Mechanism for Eliciting Probabilities," *Econometrica*, **77(2)**, 603–6.

[31] Karni, E. and Z. Safra (1995), "The impossibility of experimental elicitation of subjective probabilities", *Theory and Decision*, 38(3), 313-320.

[32] Kim, S-H. and W. Whitt (2015), "The power of alternative Kolmogorov-Smirnov tests based on transformations of the data," *ACM Trans. Model. Comput. Simul.* 25(4), article 24.

[33] Klein, WMP, J.L. Cerully and M.M. Monin (2010), "Ability, chance, and ambiguity aversion: Revisiting the competence hypothesis," *Judgment and Decision Making*, 5(3), 192–9.

[34] Klibanoff, P., M. Marinacci and S. Mukerji (2005), "A smooth model of decision making under ambiguity," *Econometrica* 73(6), pp. 1849-92.

[35] Kruger, Justin and David Dunning (1999), "Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments," *Journal of Personality and Social Psychology*, **77**, 1121–34.

[36] Li, K.K. (2011), "Preference towards control in risk taking: Control, no control, or randomize?" *Journal of Risk and Uncertainty* (43), pp. 39-63.

[37] Logg, J., J. Minson and D. Moore (2019), "Algorithm appreciation: People prefer algorithmic to human judgment," *Organizational Behavior and Human Decision Processes*, **151**, 90-103.

[38] Massoni, S., T. Gajdos and C. Vergnaud (2014), "Confidence measurement in the light of signal detection theory," *Frontiers in Psychology* 5.

[39] Massoni, S. and N. Roux (2017), "Optimal group decision: A matter of confidence calibration," *Journal of Mathematical Psychology* 79, 121-130.

[40] McKelvey, R. and T. Page (1990), "Public and Private Information: An Experimental Study of Information Pooling," *Econometrica* 58(6), pp. 1321-39.

[41] Merkle, Christoph and Martin Weber (2011), "True Overconfidence: The Inability of Rational Information Processing to Account for Apparent Overconfidence," Organizational Behavior and Human Decision Processes, 116(2), 262-71.

[42] Möbius, M., M. Niederle, P. Niehaus, and T. Rosenblat (2014), "Managing self-confidence: Theory and experimental evidence," Working Papers, No. 11-14, Federal Reserve Bank of Boston.

[43] Moore, D. (2007), "Not so above average after all: When people believe they are worse than average and its implications for theories of bias in social comparison," *Organizational Behavior and Human Decision Processes*, **102**.

[44] Owens, D., Z. Grossman and R. Fackler (2014), "The Control Premium: A Preference for Payoff Autonomy," American Economic Journal: Microeconomics, 6(4), 138-61.

[45] Razali, N.M. and Y.B. Wah (2011), "Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests," *Journal of Statistical Modeling and Analytics*, 2(1), 21-33.

[46] Ryan, R. M., V. Mims and R. Koestner (1983), "Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory," *Journal of Personality and Social Psychology*, 45(4), 736.

[47] Schattka, S. and P. Muller, (2007), "Overconfidence" in the *Encyclopedia of Social Psychology Vol. 2*, R. Baumeister and K. Vohs (eds), Sage Publications.

[48] Schlag, K., J. Tremewanz, J. van der Weele (2015), "A penny for your thoughts: A survey of methods for eliciting beliefs," *Experimental Economics*, 18:3, 457-490.

[49] Schlag, K. H. and J. van der Weele (2009), "Eliciting Probabilities, Means, Medians, Variances and Covariances without Assuming risk-neutrality", (Working Paper, Universitat Pompeu Fabra, Barcelona).

[50] Schotter, A. and I. Trevino (2014), "Belief elicitation in the laboratory," *Annual Review of Economics* (6), 103-128.

[51] Schwardmann, P. and J. van der Weele (2019), "Deception and Self-Deception," Nature Human Behaviour, July, pp 1-7.

[52] Sloof, R. and F. von Siemens (2017), "Illusion of control and the pursuit of authority," *Experimental Economics*, 20(3), 556-573.

[53] Smith, C, (1961), "Consistency in Statistical Inference and Decision," *Journal of the Royal Statistical Society Series B.* 23(1), pp. 1-37.

[54] Tversky, A. and P. Wakker (1995), "Risk Attitudes and Decision Weights, " *Econometrica* 63(6), 1255-80.