

Overconfidence and Contracts: An Experiment

Daniel Gómez-Vásquez *
Universidad del Rosario

August 2018

Abstract

We study whether overconfidence affects the choice over contracts in a real-effort task by using three different measures of overconfidence: miscalibration, illusion of control, and overplacement. We find that overplacement, which is the only measure related to the task, has a significant effect on the choice over contracts. Our design also allows us to observe overconfidence measures at different moments of time. We find that more (less) overconfident subjects remain as more (less) overconfident over time.

This paper is presented as the author's Master in Economics thesis. I thank my advisors Mariana Blanco and Çağatay Kayı for their guidance, detailed comments, and support. I am grateful for the useful comments and suggestions by Diego Aycinena, David Klinowski, Cesar Mantilla, Santiago Sautua, and the audiences of Bogotá Experimental Economics Conference - BEEC 2018 and the Bogotá Experimental and Behavioral Economics Seminar - BEBES. Finally, I thank David Ñáñez and Sergio Perilla for their support and helpful insights during the entire process. IRB approval was received from CEI - UR (Case # 2-074-CS058) before any data collection. Financial support from Universidad del Rosario's School of Economics is gratefully acknowledged.

*Master in Economics' Candidate. danielc.gomez@urosario.edu.co

1 Introduction

Overconfidence is a phenomenon in psychology of judgment that has been receiving increasing attention in behavioral and experimental economics. People tend to think that they are better than average on easy tasks but worse on hard ones, even when they are willing to pay to prove it (Williams and Gilovich, 2008). In our context, overconfidence is referred when people think their effort is more productive than it really is, i.e., they overestimate the probability of success in a determined task. There is a growing literature that studies overconfidence in different fields. For example, people tend to underestimate health risks such as risk of heart attacks, cancer, and car crashes, while they overestimate their own performance in specific tasks such as driving skills, or receiving job and wage offers.¹ Overconfidence has been also studied in the context of financial planning, bargaining, and insurance markets. These studies find that employees underestimate risk of their firm's stock, households underestimate risk of unemployment, and patients underestimate risk of death.²

Behavioral economists have models of positive thoughts and models of how decisions are made when self-image is taking into account.³ They also study business optimism over financial contracts and corporative performance.⁴ These studies find that companies directed by optimists tend to grow less, exit the industry earlier and are less profitable. Santos-Pinto (2008) and De la Rosa (2011) consider a moral hazard problem with asymmetry of beliefs. The latter model formulates the asymmetry of beliefs by overconfidence and optimism. In the equilibrium contract, there is a distinction between two effects. The first one is the *incentive* effect, which establishes that smaller increases in wages are sufficient to induce any effort level. The second one is the *wager* effect, based on the overestimation of the agent over the probability of getting a

¹See, for instance Hoch (1985), Hoorens and Buunk (1993), Kreuter and Strecher (1995), Groeger and Grande (1996), Bartl (1999), Walton and McKeown (2001), Hatakka et al. (2002), and Robb et al. (2004).

²In the context of financial planning, see Benartzi (2001). There are also the contributions by Babcock and Loewenstein (1997), Cebulla (1999), Bhattacharya et al. (2003), and Spiegler (2006).

³See Bénabou and Tirole (2002), Köszegi (2006) and Santos-Pinto and Joel (2005).

⁴See Benabou and Tirole (2003), Yildiz (2003), and Landier and Thesmar (2003).

success-contingent wage, the agent expects to receive a higher wage more often than the principal expects to pay. Sandroni and Squintani (2013) study a model of overconfidence in insurance markets. They find that effects of overconfidence depend on market structure. For instance, in monopolistic markets, overconfidence is equivalent to a change in the risk-composition of the economy.

In the literature, overconfidence can be measured in different ways. Moore and Healy (2008) present a reconciliation of three distinct ways in which the literature has defined overconfidence: excessive precision in one’s beliefs, overestimation of actual’s performance, and overplacement of one’s performance relative to others. However, there are reversals of agent’s behavior depending on the measure (Moore and Healy, 2008). Following Glaser and Weber (2007), Deaves et al. (2008), and Yang and Zhu (2016), we use three measures: miscalibration-based overconfidence (or overprecision), illusion of control (or overestimation), and overplacement.

The miscalibration-based overconfidence (MISC) (Lichtenstein et al., 1977) is the overestimation of knowledge precision, which leads to underestimation of the variance of random variables, or to excessive certainty regarding the accuracy of one’s knowledge (Moore and Healy, 2008; Menkhoff et al., 2013). MISC is measured by asking participants to construct confidence intervals of general knowledge questions. The second measure of overconfidence is the illusion of control (IOC), which is the exaggerated conviction of having control over external events (Langer, 1975). IOC is measured by using the General Self-Efficacy Scale (Schwarzer and Jerusalem, 1992). Third, overplacement (OVP) emerges when people believe themselves to be better than the others (Larrick et al., 2007). OVP is measured by eliciting a general belief about the ability of a subject relative to others (Mobius et al., 2011).

There are some experimental studies about overconfidence in different contexts. Glaser and Weber (2007) find that investors who think they are above average in terms of investment skills or past performance trade more. They also find that the miscalibration measure is not related to the measures of trading volume. Similarly, Deaves et al. (2008) find that miscalibration-based overconfidence generates additional trade in an

asset market experiment. Menkhoff et al. (2013) present the effects of overconfidence, experience, and professionalism using an online-experiment. By using non-incentivized questions they measure miscalibration, illusion of control, and unrealistic positive self-evaluations. They find that the overconfidence measures are positively but not perfectly correlated between each other. They conclude that overconfidence cannot be measured by one single measure. Yang and Zhu (2016) study the effects of overconfidence and gender on trading activity with different overconfidence measures. They use miscalibration-based overconfidence, overplacement and illusion of control. They find that the different overconfidence measures do not have correlation between each other, arguing that each measure captures only one aspect of overconfidence.

Given this context, our main question is whether the three different overconfidence measures affect the decision of the agents over contracts to perform a real-effort task. Our distinction with the previous experiments is that we use a neutral context with a real-effort task instead of the financial or trading context. As our design allows us to take some of the measures in different moments of time, we can check the stability of the measures over time, as well. This paper contributes to the literature of overconfidence and its effects on tasks, the relation between the measures and their stability over time.

The remainder of the paper is organized as follows. In Section 2, we introduce the experimental design. In Section 3, we present the main results. Section 4 concludes.

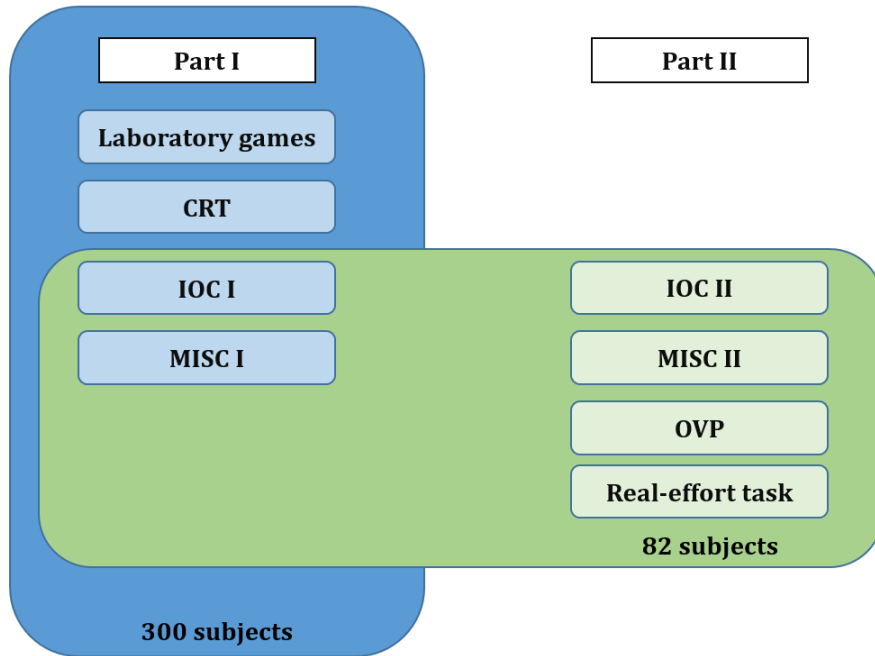
2 Experimental Design

As De la Rosa (2011) points out, under the asymmetry of beliefs (given by overconfidence) between the principal and the agent, the overconfident agent prefers a contract with different wages in case he succeeds or not over a contract that pays the same wage independent of the result. Given this hypothesis, the experiment presented here aims to prove whether overconfidence affects the propensity to choose a contingent contract over a fixed-wage contract.

Given that the literature has ambiguous results on the effect of overconfidence over

decisions, we conduct the experiment in two parts: Part I and Part II (see Figure 1). During Part I, we calculate two overconfidence measures. With these measures, we chose a sub-sample of the most and less overconfident subjects. The main objective is to have a variance on the overconfidence level of the sample for the main experiment and to find a clearer effect of overconfidence. We also measure overconfidence in Part II.

Figure 1: Experimental sessions



2.1 Part I

In Part I, we elicit two overconfidence measures: Miscalibration (MISC) and Illusion of Control (IOC). Besides, subjects play a lowest unique positive integer (LUPI) game (Östling et al., 2011) and a beauty contest game (Nagel, 1995).⁵ Then, subjects complete the cognitive reflection test (CRT) (Frederick, 2005).

The sessions from Part I have the following structure:

1. LUPI game.

⁵These games and their relation to overconfidence are for another research project.

2. Beauty contest game.
3. Cognitive Reflection Test.
4. Overconfidence (IOC).
5. Overconfidence (MISC).

MISC Measurement

Following Biaise et al. (2005), Deaves et al. (2008), and Hilton et al. (2011) for MISC measure, subjects construct 90% confidence intervals for 10 questions, where the correct answer of the question should lay in the interval. Note that a well-calibrated individual should have 90% of the correct answers contained in the reported intervals. The MISC measure is one minus the percentage of correct answers. The incentivization of the calibration test would have been injudicious. For instance, subjects asked to produce 90% confidence intervals could ensure correct calibration by making 2 narrow confidence intervals and 18 wide intervals.

Higher values of MISC measure correspond to higher levels of overconfidence. A proper calibrated individual would have a measure of 0.1. The questionnaire, translated into English, is available in Appendix B.

IOC Measurement

Following Dalton and Ghosal (2018), for IOC measure, subjects answered the questionnaire from General Self-Efficacy Scale (Schwarzer and Jerusalem, 1992) to measure generalized perceived self-efficacy, similar to the illusion of control. As the experiment was held in Colombia, we used the validated version in Spanish of the questionnaire (Manrique et al., 2005). This questionnaire consists of 10 statements. Subjects are asked to indicate how true they think each statement is about them in a scale from 1 to 4: 1 being “not true at all” and 4 being “exactly true”. This scale was validated in several studies and widely used internationally (Schwarzer and Born, 1997). In general, the scale captures the belief of performing well in a novel or difficult task. This scale

is constructed from 10 to 40, as each question is made in a 4-point scale and the final score is the sum of all 10 answers.

The higher the IOC measure, the more in control the subject appears to feel. The questionnaire is available in Appendix C.

2.2 Part II

The sessions from Part II have the following structure:

1. Practice phase.
2. Real-effort task.
3. Overconfidence (OVP).
4. Risk-aversion elicitation task.
5. Overconfidence (IOC).
6. Overconfidence (MISC).

The real-effort task is the word encryption with double randomization (WEDR), introduced in Benndorf et al. (2014). Subjects perform the task for a practice round and six periods of 2 minutes each. At the start of each period, subjects are presented with a screen as shown in Figure 2. They have to encrypt combinations of three letters (words) into three three-digit numbers. The screen consists of two rows. One displays a word to encrypt (“palabra”) and in the other one, the solution has to be entered (“código”). Below, there is an encryption table which allocates numbers to letters. The grid always displays all 26 capital letters of the Latin alphabet. Subjects have to type in the correct three-digit numbers of each letter in the boxes of the “código” row below the letter.

After all three letters are encrypted, subjects press a submit OK button. Then, they are informed whether the encryption was correctly solved. They also are given the number of correctly solved words in the current period.

This task controls for learning effect by its randomization and for some specific abilities, such as math skills. The double randomization consists in a random allocation

of numbers to all letters in the allocation table whenever subjects have encrypted a word correctly. At the same time, the positions of all letters are randomly rearranged. When subjects enter a wrong answer, they are informed by the computer. Then, the number allocations and the locations of the letters are not shuffled until subjects enter the correct answer.

The screenshot shows a software interface for a word encryption task. At the top, a status bar indicates 'Número de palabras correctas: 5' and 'Tiempo restante [sec]: 69'. The main title is 'Encriptación'. Below it, a message says 'Usted está encriptando la palabra número 6'. The word to be encrypted is 'PALABRA: Y G L', and the corresponding code boxes show 'CÓDIGO: 665 321 815'. At the bottom, a table maps letters to numbers:

Z	J	W	B	C	R	V	I	Q	G	P	M	X	D	O	L	S	A	E	H	F	Y	T	K	U	N
764	106	603	399	523	818	257	340	161	321	120	928	348	721	183	815	488	787	527	682	510	665	761	721	816	496

An 'OK' button is located in the bottom right corner.

Figure 2: Example of a problem in the WEDR task

Experimental Currency Units (ECU's) are used during the Part II sessions and exchanged into Colombian Pesos at the end of the experiment by using the rate 1:1000 COP (approximately 1 ECU: 0,35 USD). The instructions translated into English are provided in Appendix D.

Practice phase

Subjects have to solve 10 words without a time limit.

Real-effort task

The real-effort task have four stages with different payment scheme, emulating different contracts. The first two stages consist of two periods each and other two stages consist of one period each (see Figure 3):

Figure 3: Part II

No. obs						
35	Fixed-payment	Fixed-payment	Piece-payment	Piece-Payment	Bonus	Choice
47	Piece-payment	Piece-payment	Fixed-payment	Fixed-payment		
Period	1	2	3	4	5	6
Stage	1		2		3	4

1. *Fixed-payment* stage: In this stage, subjects receive a payment of 10 ECU to encrypt words in two minutes. The payment is fixed, in other words, it does not depend on the number of words that subjects encrypt. Subjects have two consecutive periods in this stage.
2. *Piece-payment* stage: In this stage, subjects receive a payment of 1 ECU for each word that they correctly encrypt in two minutes. Subjects have two consecutive periods in this stage. The performance in one of these periods is selected randomly and taken into account further in the experiment.
3. *Bonus payment* stage: In this stage, subjects receive a payment of 15 ECU if they encrypt at least one more word in this stage than in the randomly chosen period of *Piece-payment*, and 5 ECU otherwise (if subjects encrypt the same number of words or less).
4. *Choice* stage: In this stage, subjects must choose a payment scheme, among *Fixed-payment* and *Bonus payment*, for the last period. After choosing the payment scheme but before performing the task, subjects are asked the probability p that they will encrypt at least one more word in this stage than they did in the randomly chosen period of *Piece-payment* stage. The answer is used later for the measurement of risk preferences. After answering this question, each participant performs the task for 2 minutes.

In order to eliminate the order effect in the *Choice* stage, *Fixed payment* and *Piece-payment* stages have different orders in different sessions, i.e., half of the sessions started

with *Fixed-payment* followed by *Piece-payment*, while the other half started with *Piece-payment* and followed by *Fixed-payment*. The procedure is explained in Section 2.3.

OVP Measurement

We measure overplacement by following Mobius et al. (2011). Subjects are presented with two options,

1. Receive 5 ECU if their performance⁶ is among the top half of performances in the session.
2. Receive 5 ECU with probability $x \in \{0, 0.01, \dots, 0.99, 1\}$.

They are asked for which value of x they would be indifferent between the two options. Then, a random number $y \in \{0, 0.01, \dots, 0.99, 1\}$ is drawn. Subjects are paid 5 ECU with probability y if $y > x$ and otherwise received 5 ECU if their own performance was among top half of performances. To present this mechanism in a simple way, we follow Karni (2009) and Coffman (2014) and use a narrative using robots. The instructions are available in Appendix D.

The OVP measure is the value of x . The higher the OVP measure, the higher the belief of being above the median. As subjects do not know other subjects' performance or abilities, they should establish 50% as the chance of being in the first group. A metric greater than 50% refers to overconfidence.

Measurement of risk preferences

We measure risk preferences to use it as a control, and to be able to determine whether there is a relation with each overconfidence measure. Since each risk-aversion elicitation task is not suitable in this context and our focus is the change in probabilities perceived by the subjects, we make a task as similar as possible to the decision that subjects have to make in Part 3.

Thus, we offer the subjects two options:

⁶The performance is measured as the total number of words encrypted during 6 periods.

Option A: a sure payment of 10 ECU.

Option B: a lottery that pays 15 ECU with probability p and 5 ECU with probability $1 - p$.

The probability p is the value reported in the *Choice* stage by each subject.

If subjects give high values of p , it is expected from them to choose the lottery instead of the certain pay. As a second measure of risk preferences, we ask subjects a *general risk question*, to give a global assessment of their willingness to take risks in general, by answering “How willing are you to take risks, in general?” in a scale from 0 to 10 (Dohmen et al., 2011). The higher the answer, the more risk-seeking is the subject.

IOC and MISC Measurement

Both IOC and MISC measures are calculated for the subjects. For IOC measurement, subjects give answers for the general self-efficacy scale. For MISC measurement, subjects construct 90% confidence intervals for 20 questions: the same 10 questions they had in Part I and another 10 new questions. 10 new questions were asked in addition of the 10 initial questions in order to avoid contamination in the measure, as subjects may have been curious about some questions and look for the answer after the experiment. The MISC measure is calculated with all the 20 questions.

2.3 Procedures

We conducted 15 sessions: 11 sessions for Part I, with 300 subjects and 4 sessions for Part II, with 82 subjects during May 2018 at Universidad del Rosario’s Experimental and Behavioral Economics Laboratory, *REBEL*. Subjects were mainly undergraduate students from different majors, which were recruited using ORSEE (Greiner, 2015).

The sessions from Part I had an average of 27 subjects per session and no session lasted for more than one hour. In total, there were 300 subjects. During these sessions, subjects received a 10.000 COP show-up fee (approx. 3,3 USD) plus the chance of

winning a tablet.⁷ At the end of the sessions, MISC and IOC measures were calculated. Part I sessions were conducted using o-Tree (Chen et al., 2016).

In order to choose the subsample for the main experiment, the two measures (MISC and IOC) were revised and we only invited the most extreme subjects in the distribution: the most overconfident subjects and the least overconfident (or the most underconfident, if there were) subjects. This procedure decreases the internal validity of the experiment, as the results are not representative from a randomly selected participant. However, this procedure was made in order to generate variance in the subsample and to get a clearer effect of overconfidence on the choice over contracts. It is important to highlight that, by measuring MISC and IOC twice, we can check the stability of MISC and IOC.

In Part II sessions, the real-effort tasks were conducted in 4 sessions. Each one with an average around 20 subjects, for a total of 82 subjects (none of the subjects was allowed to attend more than one session, and every subject participated in the Part I). There were two sessions starting with the *Fixed-payment* stage (35 subjects). And two sessions starting with the *Piece-payment* stage (47 subjects). Part II were conducted using z-Tree (Fischbacher, 2007).

After the experiment had finished, subjects filled a post-experimental questionnaire, consisting of demographic characteristics and some questions about the subject’s perception of the experiment. Subjects received an average of 33.000 COP including the 10.000 COP show-up fee (approx. 12 USD). Part II sessions lasted for around 90 minutes.

3 Experimental results

3.1 Statistical analysis

The aim of the experiment is to analyze the effects of overconfidence on the choice over contracts. Specifically between a bonus payment and a fixed-payment contract.

⁷After all the sessions (Part I and Part II) were conducted, one of the two games, between LUPI and the beauty contest, was chosen randomly, and its winner received the tablet.

We show that there is a support to the main hypothesis: more “overconfident” subjects are more likely to choose a bonus contract over a fixed-payment contract.

Table 1: Overconfidence statistics

	Part I				Part II		
	Full Sample	Sub-sample Overconf.	Sub-sample Underconf.	Diff. <i>p</i> -value	Sub-sample Overconf.	Sub-sample Underconf.	Diff. <i>p</i> -value
MISC	0.78 (0.15)	0.86 (0.07)	0.71 (0.10)	0.000	0.81 (0.11)	0.76 (0.13)	0.082
IOC	29.24 (4.63)	33.72 (2.47)	25.84 (3.96)	0.000	34.38 (3.42)	28.70 (3.56)	0.000
OVP	- -	- -	- -		67.41 (11.24)	58.16 (14.36)	0.003
N	300	39	43		39	43	

Standard errors in parenthesis. *p*-value shown corresponds to Mann-Whitney test.

Part I

The Spearman rank correlation coefficient between the two overconfidence measures calculated in Part I (MISC and IOC) is low (0.1290) and significant at 5%. Panel A from Table 1 presents the overconfidence statistics for the two measures. Since a correct calibration in MISC is a coefficient of 0.1, almost our entire sample (97.67%) is overconfident. With the IOC measure, our sample presents a coefficient of 29.24 (in the scale from 10 to 40). Further analysis of these measures is given later, when comparing with the measures calculated in Part II sessions.

Sampling procedure

From 300 subjects, 199 were invited to Part II. First, we planned to invite subjects with the top and the bottom 100 values of each measure. However, there was no relation between these measures. We proceeded with subjects that were in the top 25% of IOC but also were in the top 50% of MISC and subjects at the bottom 25% of IOC that were also in the bottom 50% of MISC. The procedure was repeated for the top (bottom) 25% subjects in MISC but also in the top (bottom) 50% of IOC.

With this procedure, we invited subjects with lower overconfidence measures, and subjects with higher overconfidence measures. In Part I, subjects were given a unique code, announcing that the code would be used in forthcoming experiments. In Part II, we asked for the unique code in order to match the data. Table A.1 from Appendix A presents the main statistics of the measures for these subjects. In the MISC measure, the first group (more overconfident subjects) had an average measure of 0.70 while the second group (less overconfident subjects) had an average measure of 0.87. This difference is statistically different at 1% (Mann-Whitney test p -value = 0.000). For the IOC measure, more overconfident subjects had an average measure of 26.54 points while less overconfident subjects had an average measure of 33.28 points. This difference is also highly significant statistically (Mann-Whitney test p -value = 0.000). Panel B from 1 presents the statistics of the measures for the subjects from Part I that were invited to Part II and actually participated (39 subjects with higher overconfidence and 43 subjects with lower overconfidence). The difference between the groups is also statistically different at 1% for both measures (Mann-Whitney test p -value = 0.000).

Part II

Table 2 presents the Spearman rank correlation coefficients between three overconfidence measures and variables of risk preferences and ‘Choice’ (1: Bonus payment, 0: Fixed payment). While all the correlations between the overconfidence measures are positive, they are low (with coefficients between 0.0888 and 0.2084), and only the coefficient of IOC and OVP is marginally significant (at 10%). This result is consistent with the literature (Deaves et al., 2008; Glaser and Weber, 2007).

The correlation coefficients between the overconfidence measures and risk-preferences are also shown in Table 2. Risk preferences show a positive and strong correlation with IOC and OVP. Since a higher value indicates less risk aversion, greater risk-seeking is correlated with greater levels of overconfidence, especially with IOC. On the other hand, the correlations between overconfidence, risk preferences, and ‘Choice’ variable are shown. The coefficient is negative and not statistically significant for MISC and

Table 2: Correlation matrix - Part II

	MISC	IOC	OVP	Risk-taking
MISC	1.0000			
IOC	0.1118 (0.3174)	1.0000		
OVP	0.0888 (0.4275)	0.2084* (0.0603)	1.0000	
Risk-taking	0.0580 (0.6049)	0.3811*** (0.0004)	0.2491** (0.0240)	1.000
Choice	-0.1187 (0.2883)	-0.0362 (0.7470)	0.2532** (0.0217)	0.2742** (0.0127)

Obs: 82

Note: p -values in parenthesis*** $p < 0.1$, ** $p < 0.05$, * $p < 0.1$.

IOC. Hence, there is no a clear effect of these overconfidence measures over the choice between a bonus payment and a fixed-payment. However, the correlation with OVP is significant at 5% and has a positive coefficient of 0.2532. Thus, OVP has a positive effect on explaining the choice between the payment schemes. This result is also confirmed with the regression analysis. Finally, there is a positive and significant correlation between risk-preferences and ‘Choice’: the greater the risk-seeking the greater the probability to choose the bonus payment (uncertain) over a fixed-payment (certain).

Stability

In Part I, MISC was calculated by using 10 questions, while in Part II, 20 questions were asked. When using the same 10 questions, subjects were less overconfident in Part II (MISC = 0.74) than in Part I (MISC = 0.79). However, the standard deviation was bigger in Part II. This could be explained by some subjects who could searched for the answers of some questions or updated their beliefs about the general knowledge questions. When the analysis is made with the 10 ‘new’ questions, subjects were more overconfident: gave smaller intervals such that less correct answers are in the intervals reported. Thus, comparing the measures, MISC remains very close: 0.79 in Part I and 0.78 in Part II. This difference is not statistically significant (Mann-Whitney test

$p\text{-value} = 0.8770$). We also check the stability of MISC for our two sub-groups in panels B and C from Table 1. For more overconfident subjects, the measure fell from 0.86 to 0.81, i.e., the percentage of correct answers increased from 14% to 19%. For less overconfident subjects, the opposite effect was found. The measure increased from 0.71 to 0.76, i.e., the percentage of correct answers fell from 29% to 24%. However, the difference between more overconfident subjects and less overconfident subjects (0.81 and 0.76, respectively) was still significant at 10% (Mann-Whitney test $p\text{-value} = 0.0822$). It means that more overconfident subjects remained as more overconfident, while less overconfident subjects remained as less overconfident.

We continue with the stability of IOC. In both Parts I and II, IOC was calculated with the same test and 40-point scale questionnaire. Regarding subjects that participated in both parts, they were less overconfident in Part I (IOC = 29.59) than in the Part II (IOC = 31.40). The difference is statistically significant at 5% (Mann-Whitney test $p\text{-value} = 0.0355$). Panels B and C from Table 1 present the stability of IOC for our two sub-groups. IOC increased for more overconfident subjects from 33.72 in Part I to 34.38 in Part II. For less overconfident subjects, IOC increased from 25.84 to 28.70. Although the coefficients increased for both groups, the difference between them remained significant (Mann-Whitney test $p\text{-value} = 0.000$). It means, that in IOC, more overconfident subjects remained as more overconfident, while less overconfident subjects remained as less overconfident.

Regarding OVP, we find that the average subject is overconfident: the average reported probability of being above the 50% of the subjects is 62.56%. As ‘regular’ subjects (neither overconfident nor underconfident) would believe that they have a 50% probability of being above the 50% of the subjects. As shown in panels B and C from table 1, we find that less overconfident subjects (with respect to MISC and IOC in Part I) were also less overconfident with respect to OVP, with an average measure of 58.16. We found the same for more overconfident subjects, they were also more overconfident with respect to OVP, with an average measure of 67.41. This difference is statistically significant (Mann-Whitney test $p\text{-value} = 0.0031$). Table 3 presents the distribution

over two different outcomes. First is *OVP*, that refers to the probability reported by the subjects about their chances to have a performance that belongs to the top group. Second is *Real*, that refers to the fact that the subject actually had a performance above the median of their session or the performance belonged to the below-median group. From 82 subjects, 45 (shown in blue) had correct beliefs: established they had a probability greater than 50% of belonging to the best group and actually did it and subjects that established they had a low probability (lower or equal than 50%) of belonging to the best group and actually belonged to the worst group. The other 37 subjects (shown in red) had wrong beliefs about their performance. Thus, these mistaken beliefs were more common for the subjects that believed they were placed in the best group and did not.

Table 3: Overplacement statistics

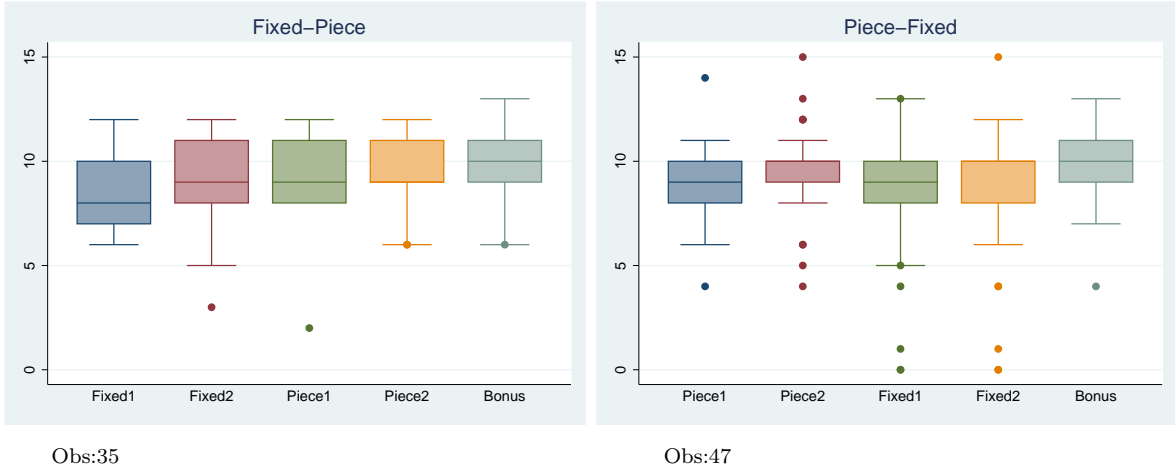
		Real		
		Above 50	Below 50	
OVP	Above 50	35	29	64
	Below 50	8	10	18
		43	39	82

Performance in the WEDR task

The analysis continues with the performance in the WEDR task. Figure 4 shows box plots for the mean number of words solved by the subjects in each of the order treatments. Each bar presents the box-whisker graph: the body represents the interquartile range (from the 25th percentile to the 75th percentile), while the whiskers go for the lower and upper adjacent values. The dots represent the outside values. The last period is not included as not all the subjects had the same payment scheme.

When comparing the graphs, it is possible to see that subjects that started with the Fixed-payment maintain a similar number of words solved in each period, while subjects that started with the Piece-payment diverged on the Fixed-payment periods

Figure 4: Mean number of words solved by period



(some subjects did not solve any word). For subjects whose session started with the Fixed payment, there were only two outside values, one for the second period of Fixed payment and another one for the first Piece-payment period. However, for subjects that started with the Piece payment, there were subjects that solved zero words in the Fixed payment periods, while other subjects solved up to 15 words in the same periods.

The performance in the task (measured with the total number of words solved in all the six periods) was not affected by the order which the subjects had during the session (Mann-Whitney test p -value=0.7891). In the analysis, we do not find any order effect for the ‘Choice’ variable (Mann-Whitney test p -value=0.1203). These results are also confirmed with the regression analysis.

Welfare analysis

The analysis finishes with the welfare analysis according to the choice between the bonus payment and fixed-payment. As Table 4 presents, from the 82 subjects, 50% (41 subjects) chose the bonus payment. According to overplacement, from the 82 subjects, 64 are overconfident. Thus, our analysis focus on these subjects. 37 from the 64 overconfident subjects (58%) chose the bonus payment, and more than half from those lose the bonus, thus, received a payment from 5 ECU instead of the 15 ECU they would have received in case of winning the bonus. It is possible to see that they would have

done better by choosing the fixed-payment instead of the bonus payment, as would received a payment of 10 ECU.

Table 4: Welfare analysis

	Overconfident			Underconfident		
	Fixed	Bonus		Fixed	Bonus	
Lose bonus	20 74.07%	20 54.05%	40 62.50%	11 78.57%	1 25.00%	12 66.67%
Won bonus	7 25.93%	17 45.95%	24 37.50%	3 21.43%	3 75.00%	6 33.33%
	27	37	64	14	4	18
	Chi-square p -value=0.102			Chi-square p -value=0.045		

This difference between loosing and wining the bonus for overconfident subjects is statistically significant at 11% (Chi-square p -value = 0.102).

3.2 Regression analysis

Choice between Fixed-payment and Bonus payment

Table 5 shows the main results from the regression analysis. *MISC* and *IOC* denote the respective measures, while *OVP* denotes a dummy variable indicating whether OVP is greater than 50 and 0 otherwise. Columns (1) and (2) show when the choice over contracts is regressed on MISC and IOC separately, neither of these measures has any explanatory power with statistically significant coefficients. However, column (3) shows that OVP has a highly significant effect on the choice of a bonus payment contract (p -value < 0.01). Column (4) shows the regression on all three overconfidence measures. Hence, OVP is the only measure to have explanatory power over the contract choice, with robust results. Note that OVP, the only overconfidence measure with a significant effect on the choice over contracts, is the only measure that is related to the task. This might illustrate the discussion of the overconfidence measures. Depending on the context, one of the different aspects of overconfidence (captured by each measure) becomes relevant.

Columns (5)-(8) allow for an exploration of additional potential determinants of the choice over contracts beyond each overconfidence measure by its own and all three together. *Risk-taking* corresponds to the risk-seeking measure calculated in Part II and is the only additional variable to have a significant effect. The higher the risk-seeking, the higher the probability of choosing a bonus (and uncertain) payment over a fixed (and certain) payment. *Order* is a dummy variable identifying sessions where the Piece-payment was played before the Fixed-payment, compared to the default session that faced the Fixed-payment before the Piece-payment. Other regressors are *Gender* (Male = 1) and *Productivity*, defined as the average words solved during the Piece-payment periods. The introduction of these control variables degrades the impact of OVP, rendering it to a lower but still significant (p -value < 0.05) coefficient. Without the controls, overconfident subjects have 36% more chance to take a bonus payment contract over a fixed payment contract, while the probability falls to 29% when including the controls in the analysis.

Table 5: Linear regression of Choice

Dep Variable: 1 [Bonus payment]	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
MISC	-0.298 (0.464)			-0.246 (0.458)	-0.463 (0.464)			-0.393 (0.459)	-0.432 (0.508)
IOC		-0.006 (0.012)		-0.008 (0.012)		-0.021 (0.015)		-0.019 (0.014)	-0.020 (0.014)
OVP			0.356*** (0.117)	0.360*** (0.116)			0.302** (0.133)	0.289** (0.134)	0.269* (0.151)
Risk-taking					0.080*** (0.030)	0.098*** (0.033)	0.063** (0.028)	0.087*** (0.032)	0.087*** (0.032)
Order					-0.159 (0.111)	-0.160 (0.111)	-0.136 (0.110)	-0.139 (0.109)	-0.138 (0.110)
CorrectBeliefs									-0.148 (0.469)
OVP×CorrectBeliefs									0.002 (0.007)
Constant	0.734* (0.369)	0.682* (0.393)	0.222** (0.099)	0.659 (0.523)	0.201 (0.586)	0.468 (0.506)	-0.145 (0.368)	0.684 (0.630)	0.745 (0.748)
Controls	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Observations	82	82	82	82	82	82	82	82	82
R-squared	0.005	0.003	0.087	0.096	0.119	0.133	0.161	0.195	0.196

Robust standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Finally, column (9) presents two more variables: *CorrectBeliefs* and the interac-

tion between this variable and the *OVP* variable. *CorrectBeliefs* is a dummy variable identifying subjects that had correct beliefs regarding their placement about their performance.⁸ It could be the case that the choice over contracts could respond to a correct prediction of performance or ability. However, these variables do not have any significant effect on the choice over contracts and do not change significantly the effect of *OVP*, only decreasing the impact of *OVP* but still significant (p -value < 0.10). With these new variables, overconfident subjects have 27% more chance to take a bonus payment contract over a fixed payment contract.

Effort in WEDR task

Table 6 presents the regression analysis for the performance in the WEDR task, measured with the total number of words correctly solved in each of the payment schemes. Columns (1)-(4) present the regression of all three measures (as in the regression of Choice, *MISC* and *IOC* denote the respective measure, while *OVP* is a dummy variable denoting overconfident subjects) to the first 5 periods. The last period is not include as not all subjects faced the same payment scheme. Columns (5)-(8) present the same regressions from the first columns, including *Order* in order to control whether the order in the staged played during the session has any effect over the number of words correctly solved.

Column (1) show that both miscalibration and overplacement have positive and significant effects on the number of words solved in the first 5 periods, while illusion of control has negative and significant effect. Column (2) presents the analysis just for the *Fixed-payment* periods. For this payment scheme, both miscalibration and overplacement have positive and significant effect, while there was no order effect for this payment scheme, as there is no significant effect of *Order* variable when it is included in the regression. Column (3) presents the analysis just for the *Piece-payment*

⁸This variable takes the value of 1 for the subjects who have correct beliefs: reported a probability greater than 50% of belonging to the best group and actually did it. Also, subjects that reported a low probability (lower or equal than 50%) of belonging to the best group and actually belonged to the worst group. The variable takes the value of 0 otherwise.

Table 6: Linear regression of performance

Dep Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of words solved	All	Fixed	Piece	Bonus	All	Fixed	Piece	Bonus
MISC	1.628** (0.784)	3.154** (1.307)	0.925 (1.169)	-0.016 (1.568)	1.640** (0.783)	3.170** (1.314)	0.947 (1.155)	-0.037 (1.587)
IOC	-0.058** (0.023)	-0.068 (0.044)	-0.059** (0.027)	-0.036 (0.037)	-0.058** (0.022)	-0.068 (0.043)	-0.059** (0.027)	-0.036 (0.036)
OVP	0.752*** (0.267)	1.008** (0.485)	0.596* (0.346)	0.553 (0.527)	0.763*** (0.270)	1.024** (0.489)	0.617* (0.350)	0.532 (0.542)
Order					0.119 (0.196)	0.174 (0.346)	0.234 (0.272)	-0.221 (0.400)
Constant	9.180*** (0.780)	7.635*** (1.302)	10.04*** (1.061)	10.55*** (1.569)	9.099*** (0.775)	7.516*** (1.287)	9.879*** (1.072)	10.70*** (1.550)
Observations	410	164	164	82	410	164	164	82
R-squared	0.044	0.067	0.043	0.024	0.044	0.069	0.048	0.028

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

periods. For this payment scheme, overplacement has positive and significant effect. While illusion of control has negative and weakly significant effect. However, when the regression is made just with this measure, there is no effect found. Finally, Column (4) presents the regression for the *Bonus payment* period: none of the measures had any significant effect on the performance at the task.

When *Order* variable is included in the analysis for each of the payment schemes (Columns (5)-(8)), the coefficients do not vary significantly from those found in the first regressions. Thus, subjects who had the *Fixed-payment* stage at the beginning of the session did not have different performance in any of the stages compared to subjects who had the *Piece-payment* stage at the beginning of the session.

Tables A.2, A.3 and A.4 present the disaggregated results for the effort in the WEDR task, according to each overconfidence measure and each payment scheme.

4 Conclusion and discussion

This paper studies the effect of overconfidence on the choice between a bonus contract (success-contingent) and a fixed-payment contract, we use three overconfidence measures. Additionally, we study the stability of the overconfidence measures in differ-

ent moments of time. First, we could confirm the hypothesis from De la Rosa (2011). Overconfident agents have higher chances of choosing an incentive contract over a fixed-wage contract. Second, it is important to highlight that the results found in this experiment are in agreement with most of the overconfidence literature; the degree of overconfidence varies with the measure, and the pairwise correlations between the measures are weakly positive. Third, from the three measures used, overplacement is the only one that explains the choice over contracts. Interestingly, it is the only measure to be related to the task. This result allows us to think that depending on the contract, a specific measure would have greater effect than other measure. Fourth, our entire sample was overconfident under the miscalibration-based overconfidence measure. As Deaves et al. (2008) found, this result may depend on the nature of the test, since it is not incentive-compatible and many subjects may not understand how to construct the intervals. However, as it does not show absolute overconfidence, the analysis can be done with relative overconfidence. Fifth, the difference between more overconfident subjects and less overconfident subjects remained significant when the measures were calculated in Part I and later in Part II. This result allows for further research, whether subject's overconfidence measures change before or after a real-effort task. Sixth, we could observe a strong relationship between risk-seeking and overconfidence measures. This allows, as well as the stability of the measures, to have further research on the relationship between risk preferences and overconfidence. Seventh, we found that more than half of the overconfident subjects that had chosen bonus contract would have done better with the fixed-payment contract. Finally, there is no clear effect of overconfidence over effort exerted in the real-effort task.

References

- Babcock, L. and Loewenstein, G. (1997). Explaining bargaining impasse: The role of self-serving biases. *The Journal of Economic Perspectives*, 11(1):109–126.
- Bartl, G. (1999). “Dan” description and analysis of post licensing measures for novice drivers. In *37. BDP-Kongress Fuer Verkehrspsychologie*.
- Bénabou, R. and Tirole, J. (2002). Self-confidence and personal motivation. *The Quarterly Journal of Economics*, 117(3):871–915.
- Benabou, R. and Tirole, J. (2003). Intrinsic and extrinsic motivation. *The Review of Economic Studies*, 70(3):489–520.
- Benartzi, S. (2001). Excessive extrapolation and the allocation of 401 (k) accounts to company stock. *The Journal of Finance*, 56(5):1747–1764.
- Benndorf, V., Rau, H. A., and Sölch, C. (2014). Minimizing learning behavior in experiments with repeated real-effort tasks. *Available at SSRN 2503029*.
- Bhattacharya, J., Goldman, D., and Sood, N. (2003). Market evidence of misperceived prices and mistaken mortality risks. Technical report, National Bureau of Economic Research.
- Biais, B., Hilton, D., Mazurier, K., and Pouget, S. (2005). Judgemental overconfidence, self-monitoring, and trading performance in an experimental financial market. *The Review of Economic Studies*, 72(2):287–312.
- Cebulla, A. (1999). A geography of insurance exclusion: perceptions of unemployment risk and actuarial risk assessment. *Area*, 31(2):111–121.
- Chen, D. L., Schonger, M., and Wickens, C. (2016). oTree: An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

- Coffman, K. B. (2014). Evidence on self-stereotyping and the contribution of ideas. *The Quarterly Journal of Economics*, 129(4):1625–1660.
- Dalton, P. S. and Ghosal, S. (2018). Self-confidence, overconfidence and prenatal testosterone exposure: Evidence from the lab. *Frontiers in Behavioral Neuroscience*, 12:5.
- De la Rosa, L. E. (2011). Overconfidence and moral hazard. *Games and Economic Behavior*, 73(2):429–451.
- Deaves, R., Lüders, E., and Luo, G. Y. (2008). An experimental test of the impact of overconfidence and gender on trading activity. *Review of Finance*, 13(3):555–575.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.
- Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19(4):25–42.
- Glaser, M. and Weber, M. (2007). Overconfidence and trading volume. *The Geneva Risk and Insurance Review*, 32(1):1–36.
- Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with orsee. *Journal of the Economic Science Association*, 1(1):114–125.
- Groeger, J. A. and Grande, G. E. (1996). Self-preserving assessments of skill? *British Journal of Psychology*, 87(1):61–79.
- Hatakka, M., Keskinen, E., Gregersen, N.-P., Glad, A., and Hernetkoski, K. (2002). From control of the vehicle to personal self-control; broadening the perspectives to driver education. *Transportation Research Part F: Traffic Psychology and Behaviour*, 5(3):201–215.

- Hilton, D., Regner, I., Cabantous, L., Charalambides, L., and Vautier, S. (2011). Do positive illusions predict overconfidence in judgment? A test using interval production and probability evaluation measures of miscalibration. *Journal of Behavioral Decision Making*, 24(2):117–139.
- Hoch, S. J. (1985). Counterfactual reasoning and accuracy in predicting personal events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(4):719.
- Hoorens, V. and Buunk, B. P. (1993). Social comparison of health risks: Locus of control, the person-positivity bias, and unrealistic optimism. *Journal of Applied Social Psychology*, 23(4):291–302.
- Karni, E. (2009). A mechanism for eliciting probabilities. *Econometrica*, 77(2):603–606.
- Köszegi, B. (2006). Ego utility, overconfidence, and task choice. *Journal of the European Economic Association*, 4(4):673–707.
- Kreuter, M. W. and Strecher, V. J. (1995). Changing inaccurate perceptions of health risk: results from a randomized trial. *Health psychology*, 14(1):56.
- Landier, A. and Thesmar, D. (2003). Financial contracting with optimistic entrepreneurs: Theory and evidence. *Available at SSRN 479581*.
- Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, 32(2):311.
- Larrick, R. P., Burson, K. A., and Soll, J. B. (2007). Social comparison and confidence: When thinking youre better than average predicts overconfidence (and when it does not). *Organizational Behavior and Human Decision Processes*, 102(1):76–94.
- Lichtenstein, S., Fischhoff, B., and Phillips, L. D. (1977). Calibration of probabilities: The state of the art. In *Decision Making and Change in Human Affairs*, pages 275–324. Springer, Dordrecht.

- Manrique, G., Chia, S., and Valdez, P. (2005). Género, ansiedad social, autoeficacia general y autoeficacia en situaciones sociales. *Revista de Psiquiatria y Salud Mental Hermilio Valdizan*, 6:43–54.
- Menkhoff, L., Schmeling, M., and Schmidt, U. (2013). Overconfidence, experience, and professionalism: An experimental study. *Journal of Economic Behavior and Organization*, 86:92–101.
- Mobius, M. M., Niederle, M., Niehaus, P., and Rosenblat, T. S. (2011). Managing self-confidence: Theory and experimental evidence. Technical report, National Bureau of Economic Research.
- Moore, D. A. and Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, 115(2):502.
- Nagel, R. (1995). Unraveling in guessing games: An experimental study. *American Economic Review*, 85(5):1313–1326.
- Östling, R., Wang, J. T., Chou, E. Y., and Camerer, C. F. (2011). Testing game theory in the field: Swedish lupi lottery games. *American Economic Journal: Microeconomics*, 3(3):1–33.
- Robb, K. A., Miles, A., and Wardle, J. (2004). Subjective and objective risk of colorectal cancer (UK). *Cancer Causes and Control*, 15(1):21–25.
- Sandroni, A. and Squintani, F. (2013). Overconfidence and asymmetric information: The case of insurance. *Journal of Economic Behavior and Organization*, 93:149–165.
- Santos-Pinto, L. (2008). Positive self-image and incentives in organisations. *Economic Journal*, 118(531):1315–1332.
- Santos-Pinto, L. and Joel, S. (2005). A model of positive self-image in subjective assessments. *American Economic Review*, 95(5):1386–1402.

- Schwarzer, R. and Born, A. (1997). Optimistic self-beliefs: Assessment of general perceived self-efficacy in thirteen cultures. *World Psychology*, 3(1-2):177–190.
- Schwarzer, R. and Jerusalem, M. (1992). Self-efficacy as a resource factor in stress appraisal processes. *Self-efficacy: Thought control of action*, 195213.
- Spiegler, R. (2006). Competition over agents with boundedly rational expectations. *Theoretical Economics*, 1(2):207–231.
- Walton, D. and McKeown, P. (2001). Drivers biased perceptions of speed and safety campaign messages. *Accident Analysis and Prevention*, 33(5):629–640.
- Williams, E. F. and Gilovich, T. (2008). Do people really believe they are above average? *Journal of Experimental Social Psychology*, 44(4):1121–1128.
- Yang, X. and Zhu, L. (2016). Ambiguity vs risk: An experimental study of overconfidence, gender and trading activity. *Journal of Behavioral and Experimental Finance*, 9:125–131.
- Yildiz, M. (2003). Bargaining without a common prioran immediate agreement theorem. *Econometrica*, 71(3):793–811.

Appendix

A Additional tables

Table A.1: Sampling procedure statistics

Measure	Type	Obs.	Mean	SD
MISC	Less overconfident	112	0.70	0.12
	More overconfident	87	0.87	0.08
IOC	Less overconfident	112	26.54	3.47
	More overconfident	87	33.28	2.43

Table A.2: Linear regression of performance, Fixed-payment periods

Dep Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of words solved								
MISC	2.910** (1.310)			3.154** (1.307)	2.922** (1.319)			3.170** (1.314)
IOC		-0.0563 (0.0436)		-0.0684 (0.0439)		-0.0559 (0.0429)		-0.0677 (0.0433)
OVP			0.925* (0.497)	1.008** (0.485)			0.942* (0.501)	1.024** (0.489)
Order					0.129 (0.351)	0.0924 (0.343)	0.174 (0.352)	0.174 (0.346)
Constant	6.466*** (1.104)	10.52*** (1.296)	8.028*** (0.457)	7.635*** (1.302)	6.402*** (1.181)	10.47*** (1.217)	7.941*** (0.531)	7.516*** (1.287)
Observations	164	164	164	164	164	164	164	164
R-squared	0.023	0.011	0.026	0.067	0.024	0.012	0.028	0.069

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table A.3: Linear regression of performance, Piece-payment periods

Dep Variable: Number of words solved	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MISC	0.741 (1.187)			0.925 (1.169)	0.761 (1.174)			0.947 (1.155)
IOC		-0.0538** (0.0254)		-0.0594** (0.0266)		-0.0530** (0.0258)		-0.0585** (0.0270)
OVP			0.540 (0.355)	0.596* (0.346)			0.563 (0.360)	0.617* (0.350)
Order					0.214 (0.271)	0.190 (0.271)	0.246 (0.274)	0.234 (0.272)
Constant	8.784*** (0.947)	11.05*** (0.778)	8.944*** (0.323)	10.04*** (1.061)	8.678*** (0.922)	10.95*** (0.821)	8.821*** (0.370)	9.879*** (1.072)
Observations	164	164	164	164	164	164	164	164
R-squared	0.003	0.019	0.017	0.043	0.007	0.022	0.022	0.048
Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1								

Table A.4: Linear regression of performance, Bonus payment period

Dep Variable: Number of words solved	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MISC	-0.145 (1.596)			-0.0156 (1.568)	-0.168 (1.616)			-0.0365 (1.587)
IOC		-0.0319 (0.0353)		-0.0356 (0.0366)		-0.0330 (0.0349)		-0.0364 (0.0363)
OVP			0.524 (0.543)	0.553 (0.527)			0.505 (0.558)	0.532 (0.542)
Order					-0.244 (0.390)	-0.255 (0.390)	-0.210 (0.400)	-0.221 (0.400)
Constant	9.968*** (1.259)	10.85*** (1.132)	9.444*** (0.504)	10.55*** (1.569)	10.09*** (1.236)	11.00*** (1.121)	9.549*** (0.579)	10.70*** (1.550)
Observations	82	82	82	82	82	82	82	82
R-squared	0.000	0.007	0.016	0.024	0.005	0.012	0.019	0.028
Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1								

B Questions used in the Miscalibration measure

Questions used in both Part I and Part II:

	Lower bound	Higher bound
Martin Luther King's age at death		
Gestation period (in days) of an Asian elephant		
Number of countries that are member of the OPEC		
Percentage of total area in the world covered by water		
Year in which Bell patented the telephone		
Length of the Nile River (in meters)		
Air distance from London to Tokyo		
Number of books in the Old Testament		
Weight of an empty Boeing 747		
Number of babies born in world in 2001 (per 1000 people)		

'New' questions used only in Part II:

	Lower bound	Higher bound
Year in which J. S. Bach was born		
Year in which Newton discovered universal gravitation		
Year in which Mozart wrote his first symphony		
Number of medals that Greece won at the first Olympic Summer Games in 1896		
Deepest known point in the oceans (in kmts.)		
Height of Sears Tower in Chicago (in mts.)		
Number of joints in the human body		
World-wide life expectancy at birth in 2001		
Age of sun in millions of years		
Diameter of the moon (in kmt.s)		

C Self-Efficacy Scale (Schwarzer and Jerusalem, 1992)

1. ... I can always manage to solve difficult problems if I try hard enough.
2. ... If someone opposes me, I can find the means and ways to get what I want.
3. ... It is easy for me to stick to my aims and accomplish my goals.
4. ... I am confident that I could deal efficiently with unexpected events.
5. ... Thanks to my resourcefulness, I know how to handle unforeseen situations.
6. ... I can solve most problems if I invest the necessary effort.
7. ... I can remain calm when facing difficulties because I can rely on my coping abilities.
8. ... When I am confronted with a problem, I can usually find several solutions.
9. ... If I am in trouble, I can usually think of a solution.
10. ... I can usually handle whatever comes my way.

D Instructions (Translated from Spanish)

General Structure of the Experiment

This is a experiment that consists of a practice period and four sections. Before each section, you will receive specific instructions for the section. This are the practice period instructions.

During the experiment you have to do a task. The task consists of **encrypt** combinations of letters in numbers. In the task, three letters yield a “word”. You have to allocate the three digit number that corresponds to each letter. The equivalence table between letters and numbers can be found in the lower part of the screen. The letters to encrypt appear in the middle and the encryption boxes below each letter. For that purpose, please consider the following image:

Número de palabras correctas: 5
Tiempo restante [sec]: 69

Encriptación

Usted está encriptando la palabra número 6

PALABRA: Y G L
CÓDIGO: 665 321 815

Z	J	W	B	C	R	V	I	Q	G	P	M	X	D	O	L	S	A	E	H	F	Y	T	K	U	N
784	106	603	399	523	818	257	340	161	321	120	928	348	721	183	815	488	787	527	682	510	665	761	721	816	496

OK

In the example, the subject has encrypted five words correctly (see the superior horizontal bar). Here, the three letters “Y”, “G” and “L” have to be encrypted. The solution comes from the equivalence table:

- For “Y” applies 665.
- For “G” applies 321.
- For “L” applies 815.

Once you have registered the three numbers, click “OK”.

The computer checks that all the letters have been encrypted correctly. Then a new “word” (also consisting of three letters) appears in your screen to be encrypted. The words that you have to encrypt are chosen by the computer randomly.

Take into account that the equivalence table changes after every “word” encrypted correctly. The new equivalence table is generated in two steps:

1. The computer selects randomly in the table a new set of three-digit numbers that are going to be used to encrypt the letters.
2. The computer changes the position of the letters in the equivalence table. Note that the computer uses all the 26 letters of the alphabet.

Note that if a new “word” appears, you have to click in the first of the three boxes. Otherwise no input is possible. The computer will mark in red font the wrong inputs after click the

“OK” button. Bear in mind that if you make a wrong input:

- The word to encrypt will not change until you encrypt it correctly.
- The computer will erase all the entries in the word, including the letters that have encrypted correctly.
- The equivalence table remains unaltered, that is, the allocation of numbers to each letter remains the same. The position of the letters does not change either.

Practice Period

- The experiment starts with a Practice period where you have to encrypt exactly 10 words.
- Bear in mind: correct solutions in the Practice period will not generate payments.
- The general idea of this period is make you as familiar as possible to the task before the periods that will count to your payment start.

Instructions (Section 1)

The practice period just finished. Now starts the first section of the experiment. In this section, you have to do the same task: encrypt “words” to numbers.

This section is composed of four parts. The two first parts are composed by two periods and the two last parts are composed by one period. Each one of the periods will have an exact duration of two minutes. When each period finishes, the computer will count only the “words” that have been encrypted correctly during the two minutes of this period.

To determine your payments, the computer will select one of the six period randomly to pay you, where each period has the same chances of being chosen. As you don’t know with anticipation which period will be chosen, consider each one as if it was the one that determines your payments. In total there are four parts, each one corresponds to a payment scheme, these will be described next:

Part 1 - Fixed payment

This scheme consists of a fixed payment of 10 ECU. As this payment is fixed, does not depend on the number of “words” that you encrypt correctly.

- This part will have **two** periods.
- At the end of each period, you will receive information in the screen about the number of “words” encrypted correctly in the period.

If at the end of the experiment one of these periods is selected randomly to determine your payment, you will receive 10 ECU without taking into account the number of “words” encrypted correctly. Remember that only one of the six periods will be selected to determine your payment.

Part 2 - Payment per word

This scheme consists in a payment that depends only on the “words” that have been encrypted correctly. You will receive 1 ECU per each correct “word”.

- This part will have **two** periods.
- At the end of each period, you will receive information in the screen about the number of “words” encrypted correctly in the period.

If at the end of the experiment one of these periods is selected randomly to determine your payment, you will receive your earnings according to the number of “words” encrypted correctly in that period. Remember that only one of the six periods will be selected to determine your payment.

Part 3 - Bonus payment

In this scheme your payment depends on the fulfillment of an objective. The objective is to encrypt at least one more word in this period than in one of the two periods of payment per word. The period that sets your objective will be chosen randomly by the computer and will only be known at the end of the experiment if this period is chosen as the one that determines your payment.

- You will receive 15 ECU if you can encrypt at least one more word in this period than in the **payment per word** period chosen randomly to set your objective.
- You will receive 5 ECU in case you encrypt the same number of “words” or less.

- At the end of the period, you will receive information in the screen about the number of “words” encrypted correctly in the period.

Part 4 - Choice

In this period you have to do once more the encryption task. At the beginning of the period you have the chance of choosing the payment scheme that you prefer. Your options are **fixed payment** or **bonus payment**. After choosing the payment scheme, you do the task for one unique period.

If at the end of the experiment this period es chosen randomly to determine your payment, the earnings of this period will correspond to your choice (**fixed payment** or **bonus payment**) and your performance (in the case of **bonus payment**). Remember that only one of the six periods will be selected to determine your payment.

Instructions (Section 2)

In this section we will divide the session in two groups according the performance in the encryption task (the total number of “words” encrypted correctly in the six periods). Each group will have half the subjects of the session. The first group has the 50% of the subjects with the higher scores and the second group the 50% with the lower scores.

In this section, we will have robots that could do the encryption task. There are 100 different robots, with numbers from 1 to 100, where the number corresponds to the efficiency of the robot. In this case, we measure the efficiency with the probability (in percentage) that the robot makes the encryption task for the six periods and has a performance that places him in the first group. That is, Robot #1 has a probability of 1% that his performance will place him in the first group and a probability of 99% that his performance will place him in the second group, Robot #2 has a probability of 2% that his performance will place him in the first group and a probability of 98% that his performance will place him in the second group, ..., and so on until Robot #100, that has a probability of 100% that his performance will place him in the first group. For example, Robot #75 will be placed in the first group with a probability of 75% and will be placed in the second group with a probability of 25%.

Now you have to decide which robots you allow to make the task instead of you. You have

to make a decision for each one of the 100 robots: You prefer to use the robot's performance or your own?

You will write how sure you feel that your performance will place you in the first group. You will do that by choosing which robots you allow to make the task instead of you to define if you are in the group with the best performance or in the worst one. For example, you have more chances of being placed in the first group by using your own performance instead of the performance of Robot #1. The same happens if you use Robot#100 performance instead of your own performance.

Your task is resumed in deciding, for each robot, if you prefer that he replaces you in the task, or if you prefer to do the task yourself. As the robots are ordered from #1 (the most inefficient of the robots, the one which will loose against you for sure) to the #100 (the most efficient of the robots, the one that will win against you), we just need you to tell us from which robot you prefer that your performance be replaced by the robot's performance.

Hence, you must complete the phrase "I think my performance has a probability of ___% of belonging to the first group. The robot should do the task for me if its efficiency is greater or equal than that". You have to write a number between 1 and 100.

The computer will choose a robot randomly. Each robot has the same chances of being selected. If the chosen robot has an efficiency greater or equal to the probability that you chose, the robot will do the task and its performance will be taken instead of yours. If the chosen robot has an efficiency lower than the probability you chose, your own performance will be taken instead of the robot's performance.

For example, if you choose 75 as your probability of being in the first group, and the Robot#90 is chosen randomly, the robot's performance will be taken instead of yours. This robot has a probability of 90% of having a performance that will place him in the first group. If you choose 75 as your probability of being in the first group, and the Robot#20 is chosen randomly, your own performance will be taken instead of the robot's performance.

You will not know which robot was chosen neither its performance until the end of the experiment. For the payment of this section, you will receive 5 ECU if your performance belongs to the first group, no matter if is the robot's performance or your own. Take into account that the result of this section will not change your payments obtained in the encryption task.

Instructions (Section 3)

In this section you have two options: a safe payment and a lottery, you have to indicate which one of the two options you prefer. If you choose the safe payment, your payment will be 10 ECU and if your choice is the lottery, the computer will make it and inform you about the result.

Example:

Option A 10 ECU

Option B 15 ECU with 70% probability
5 ECU with 30% probability

If you choose option A, you will receive the 10 ECU. If you choose option B, the lottery will be made. In this case, with a probability of 70% you will receive 15 ECU and with the 30% remaining, you will receive 5 ECU.

Instructions (Section 4)

Survey 1

Next you will see some statements that are related with ways people face different situations of life. Read carefully each statement and indicate how true you think each statement is for you.

(1 = not true at all, 2 = hardly true, 3 = moderately true, 4 = exactly true)

Select your answer at the end of each statement.

Survey 2

In the past part of the experiment, you will have 20 questions, and for each one you have to think of an interval in which you are 90% sure that the correct answer lies in the interval. This interval is composed of a lower bound and a higher bound. We want you to give a range in which you are 90% sure that the correct answer lies inside these limits. That is, from 10 answers that lie in your mind as correct answer, you must drop one, in one of the two extremes.

Take into account that if your interval (the difference between the higher bound and the lower bound) is too wide, the correct answer will lie in your interval more than 90% of the

times.

Example:

	Lower bound	Higher bound
Colombia's population in 2016 (in millions)	30	90

If your interval is too narrow, the correct answer will lie in your interval less than 90% of the times.

Example:

	Lower bound	Higher bound
Colombia's population in 2016 (in millions)	48.5	48.75

If you are sure of any question, the higher bound and the lower bound would be the same.

Example:

	Lower bound	Higher bound
Number of months the year has	12	12