# Lying to be Fair

Gönül Doğan[1] and Rikstje Anneke Roggema[2]

**Abstract**

One of the main arguments people use for cheating is reciprocity. Cheating becomes a tool for establishing fairness when others cheat. We experimentally investigate whether there is reciprocal lying. We vary who can lie and the type of payoff scheme. In a real effort task, we find that varying who lies does not affect lying rates in any of the payoff schemes we employ. We further study whether lying depends on ability and fairness considerations. We observe two effects independent of lying: own ability correlates with beliefs about the others' ability, and higher ability people consider unequal outcomes fairer compared to low ability people.

**Keywords: cheating, reciprocity, competition, real effort**
**JEL codes: C92, D63, J31**

[1] Corresponding Author. Department of Economics, University of Amsterdam, 1018WB, Amsterdam, the Netherlands. gonul2517@gmail.com, +31205254229.
[2] Department of Economics, University of Amsterdam, 1018WB, Amsterdam, the Netherlands. Current Address: Rocket Internet, I 22 Pan Hlaing Road, Sanchaung Township, Yangon, Myanmar. r.a.roggema@gmail.com, +959254374825.

## 1. Introduction

In many real-life situations, one's honesty depends on whether others are honest. Consider tax evasion. In countries with high tax evasion levels, people mention government corruption as one of the most important reasons that justify tax evasion. Tax compliance correlates positively with the strength of the perceived social norm of tax compliance. Also, acceptance attitudes towards tax evasion correlate with the number of tax evaders a person knows (see e.g. Wallschutzky, 1984; Becker et al. 1987; Wenzel, 2004). These findings point to the fact that cheating on taxes is easier when compliance norm is broken.

Cheating can also be a tool for establishing the 'fair' outcome; that is, the outcome that would have been achieved if everyone were to be honest. Continuing with the tax example, consider the effect of non-compliance by a large group of people: Because the burden created by non-compliers is substantial[3], tax evasion can be seen as a way of off-setting the injustice done by other tax evaders. Therefore, people who would pay their taxes honestly if the majority were to be honest might cheat on their taxes when others cheat. Similar arguments are routinely voiced in professional sports. When Lance Armstrong was convicted of cheating, his main line of defence was that everyone else was doing it. In an interview, he said that he did not view doping as cheating, but rather, as a "level playing field" (Telegraph Sport, 2013). Since everyone dopes, the only way a cyclist has a chance of winning is by doping. Yet another example is summarized in this quote (see Moore, 2013) by one of the protestors when students at a school in China were not allowed to cheat in the university entrance "gaokao" exam: "We want fairness. There is no fairness if you do not let us cheat."

In this paper, we experimentally investigate whether cheating is used as a tool to restore fairness. We employ a design that separates cheating for selfish monetary gains from reciprocal cheating. We introduce a real effort task and give some people the opportunity to lie over their performance in the task. Participants are matched in a pair. To find out whether cheating takes place to restore equity, we vary who can lie: In the "one-party" treatments only one person in the pair can lie, and in the "two-party" treatments both parties can lie. If participants expect their counterpart to lie, then they might lie to restore fairness in the two-party treatments but such a

[3] Such a burden can indeed be very large; the estimated effect of tax evasion on income inequality in 2005 in Greece was a 9.7 percent change in the Theil measure of the income distribution implying huge costs on honest tax-payers (Matsaganis and Flevotomou, 2010).

reasoning does not apply in the one-party treatments. Thus, if reciprocity does indeed play a role in lying, then we should see less lying in the one-party treatments than in the two-party treatments. This is obviously conditional on the fact that subjects expect lying by others; if there is no expectation of lying, then reciprocity cannot play a role[4].

For our investigation, we implement three payoff schemes that vary the effect of lying on the final outcome. These are the piece-rate, tournament, and pie-sharing. In the piece-rate, the only effect of one's lying on the other is via the weakening of the social norm of honesty. Thus, if we observe any effect of varying who can lie in the piece-rate, it would be a clear demonstration that people care about others' (possible) lies even when those lies do not hurt them. In the tournament, on the other hand, cheating might take place to restore the fair outcome. When a person thinks that she performs well and should win the tournament, but also expects the other to cheat and win instead, lying is the only way of winning. This was the reasoning of Lance Armstrong for cheating in the tournaments he competed in. Finally, we introduce the pie-sharing payoff in which parties get a share of a fixed-sized pie according to the ratio of their declarations. Pie-sharing is representative of situations in which funds are distributed on the basis of performance within a group of candidates. Bonuses distributed according to relative performance in organizations or government funds distributed among universities, hospitals or schools fit into this category. Since the pie is fixed, inflating one's performance leads to getting a bigger share of the pie at a cost to the other. Examples of such deception are plenty. For instance, in many schools in the U.S. with a high number of lower performing students, teachers or administrators misrepresent the academic performance of their pupils to secure more government funding for their school (for a recent case, see Banchero, 2014). Importantly however, the harm caused by cheating under the pie-sharing payoff scheme is not as much as in the tournament when lying changes the winner. Therefore, we would expect pie-sharing to be less sensitive to who can lie than tournament.

The examples above suggest that lying in tournaments might be driven by high performers who have a lot to lose whereas pie-sharing creates incentives for lying for low performers because they gain relatively more by lying. Fairness concerns, other than aiming to restore the honest outcome, might also play a role in lying behaviour. A natural candidate is equal division of resources; if equality of payoffs is considered to be the fair outcome, then lower ability people might lie more

[4] Previous studies show substantial amount of lying or cheating using different games; ex-ante we do not have any reason to expect otherwise.

in competitive environments. To investigate these relationships further, we elicit subjects' beliefs about the performance of others (incentivised), and their fairness considerations via a questionnaire administered after the experiment. We posit that there is an underlying relationship between a person's own ability and her estimation of others' ability. We can find out this relationship by studying the answers of subjects who cannot lie. If lying interacts with real ability, we would see a different relationship between a person's own ability and her estimation of others' ability when subjects can lie.[5] Similarly, if we see that fairness considerations explain declared performances differently than real performances, we can infer that they play a role in lying behaviour.

We find that overall, few people lie, and maximal lying is rare. Average amount of lies are highest in the piece-rate and there is no evidence for lying in the tournament. Further, we find no evidence for reciprocal lying in our experiments; that is, the amount of lying is not significantly higher when both parties in a pair can lie compared to when only one of them can lie. We argue that this might be a result of low lying rates in our experiments. In the analytical task, women perform worse than men and lie more, but only in the piece-rate. There is no evidence that women lie in the tournament or pie-sharing. Men seem unresponsive to the payoff scheme.

Finally, our data suggests that lying behaviour is not related to ability, nor beliefs on fairness. As expected, there is a relationship between ability and the estimated ability distribution: Lower ability people underestimate the ability of others whereas higher ability people overestimate. However, this relationship does not vary with lying. Our analysis also shows that fairness considerations are negatively correlated with one's ability: the higher the ability of a person, the more that person thinks it is fair to distribute money unequally. However, there is no indication that such fairness concerns are reflected in lying behaviour.

## 2. Literature

[5] We know the real performance of a large group of subjects, and this is our control group. We, however, do not know the real performance of subjects who declare their performance, and therefore we do not know whether and how much they lie. Thus, we can only indirectly infer the relationship between ability and lying. Further, we decided not to elicit the beliefs of subjects' about others' lies to make sure that subjects who did not realize the possibility of lying remained ignorant. This way, we could ensure that future sessions would not be affected. Given our results, this seems to be the right choice.

This paper contributes specifically to the literatures on deception and gender differences. We know from the growing experimental literature on cheating that people do have a preference for honesty, even at a cost to themselves. For example, Gneezy (2005) studied the effect of the absolute and relative consequences of lies on the participants' propensity to lie. In a two player cheap-talk sender-receiver game, he varied the size of the lie by varying the gains of both the sender and the receiver. He found that people cared about the cost of their lies. Another common finding is that most people lie a little rather than maximally in many different type of experiments (see e.g., Erat and Gneezy, 2012; Fischbacher, and Heusi, 2008; Gibson, Tanner, and Wagner, 2013; Gneezy, Rockenbach, and Serra-Garcia, 2013; Hurkens and Kartik, 2009; Lundquist et al., 2009)

Recent work on the effect of competition on cheating behaviour provides mixed results. Whereas some studies find higher levels of cheating with competition for status or money (Belot and Schröder, 2013; Pascual-Ezama, Prelec, and Dunfield, 2013) others report no change (Schwieren and Weichselbaumer, 2010). For example, Schwieren and Weichselbaumer (2010) conducted a computerized maze solving game comparing cheating behaviour in tournament and piece-rate payoff schemes. They found that about 40 percent of subjects cheated irrespective of the payoff scheme. Women cheated more under tournament and men less, but when performance was taken into account, the gender differences disappeared. In Pascual-Ezama, Prelec, and Dunfield (2013), subjects were paid for finding 10 instances of two consecutive letters on a sheet with a seemingly random sequence of letters. They first replicated the Ariely, Kamenica and Prelec (2008) study: Participants were paid a piece-rate for every handed in sheet, and how these handed-in sheets were checked were varied. In one condition all sheets were checked by the experimenters, in the second condition not checked, and finally in a third condition, the sheets were passed through a shredder. They found substantial cheating when the sheets were not checked. They further employed a social competition (announcement of the winner to other subjects) and economic competition (additional money for being a winner) treatments. They found more cheating under both of the competition treatments. Their design, however, does not allow for an estimation of percentage of cheaters, nor the analysis of gender.

Finally, sabotage that involves deception is also relevant for our work. Carpenter et al. (2010) studied sabotage within a real effort experiment: they asked subjects to prepare letters and envelopes, and they employed piece-rate and tournament schemes. They found that subjects deliberately misrepresented the quality and quantity of the competitor's work in the tournament regardless of gender. Likewise, Charness, Masclet, and Villeval (2013) and Dato, and Nieken, (2014) find more sabotage via cheating tournaments. Finally, Rigdon and D'Esterre (2012) let their subjects inflate their own performance and also deflate the other participant's performance. They found that people inflated their own performance to some extent, but they were not willing to sabotage the work of someone else. They did not find an effect of competition for either type of cheating behaviour.

The mixed results of the effect of competition on cheating might stem from different expectations of what others would do in different games. High levels of cheating might be more likely when the possibility of cheating is obvious, and subjects' expectation of cheating is high. In this paper, our main contribution to this literature is to study whether there is "reciprocal" cheating. To our knowledge, this is the first study in the literature to investigate the possibility of cheating as a fairness tool.

## 3. Experimental design

### Procedures

The experiments were conducted at the Center for Research in Experimental Economics and political Decision Making (CREED) of the University of Amsterdam. Subjects were recruited via the online recruitment system of CREED and were mostly undergraduates from a wide variety of majors. Each subject could participate in only one session, and all treatments were across subjects, i.e. in each session only one treatment was run. There were 7 treatments conducted in 24 sessions with 472 students from different disciplines. The experiments lasted about one hour including the time spent on payment. Average pay was 10,7 euros including 3 euros show-up fee.

At the beginning of the experiment, instructions were read out loud. In each session, participants were randomly assigned to one of three player labels: Player A, B or C. Subjects were told that there are two parts in the experiment, and their payment in the experiment is based on the task that they do in the first part. Further, they were informed that the determination of their payment is conducted in the second part. The task involved 14 questions, and all subjects were given 10 minutes to write down their answers on an answer sheet[6]. A and B players were randomly matched for the payment of the task, and the C players corrected the answer sheets. We employed the correctors from among the subjects to make it clear that we did not correct their answer sheets. A and B players were told that the task forms the basis of their payment and that they should give as many correct answers as possible. Since subjects were not told about the payoff scheme until after they finish answering the questions, we do not expect to have any treatment effects on the real performance of our subjects.

After finishing the task, all player As and Bs were instructed to put their answer sheets blank page facing up. An experimenter collected the answer sheets without looking at them. Further, there were no identifiers on the answer sheets; we kept track of which answer sheet belongs to which table via the order of collection. The answer sheets were given to the randomly assigned C players. Player C's were instructed to highlight the correct answers with a highlighter. We made sure that the answer sheets were corrected in exactly the same order. The payment of C players were done by randomly picking one of the corrected answer sheets, and controlling whether the correction is fully correct. If the highlighted answers were correct, the C player earned 10 euros, otherwise nothing. C players could take as much time as they needed to make the corrections. After C players finished correcting the answer sheets, the answer sheets were distributed back to A and B players, again blank page facing up. We explained this procedure in detail in the instructions.

After the A and B players received their corrected answer sheets, the instructions for the second part of the experiment were distributed and then read out loud. The second part determines the payment of A and B players. First the payoff scheme was explained. Then subjects were told who will receive a declaration form (only A players in the one-party treatments or both A and B players in the two-party treatments), and that the declaration forms will be used by an experimenter, who was not involved in the running of the experiment, to calculate their payment. Similar to the first part, all forms were collected blank page facing up. Finally, the answer sheets of all players were handed back to C players, and C players highlighted all the answers. This final step was to minimize the risk of cheating across sessions by using the right answers from subjects of previous sessions.

[6] We employed a real effort task because whether money is earned by putting in effort or by a random device might make a difference when cheating. Previous experiments suggest that earned money is treated differently than money given by the experimenter. See for example the results by Oxoby and Spraggon (2008).

Finally, the declaration forms and the answer sheets were returned to the subjects. Subjects were told that they can keep the declaration forms and the answer sheets, and most of them did so.

After the experiment is finished, subjects were asked to fill in a questionnaire stating their gender, studies, the number of experiments they participated in that academic year, the number of times they took a GRE/GMAT type of test, their beliefs on the distribution of correct answers, their guess of the average correct answer, and how they think some money should be divided between a pair under different combinations of correct answers. To elicit the A and B players' beliefs on the distribution of correct answers, they were told that a certain number of people have done the task before[7], and they were asked to guess how many of those subjects have answered 0, 1, 2, … 14 correct answers. They could earn an additional 6 Euros if their guesses matched that of the distribution of the no-lie condition[8], and otherwise every difference cost them 50 Eurocents. Finally, we did not use the word lying anywhere in our experiment, nor explicitly suggested the possibility of cheating.

**Payment Structures**

We implemented three different payoff structures that vary the effect of lying. These payoff structures are depicted across the three rows of Table 1. To study the effect of reciprocity in lying, we varied who can lie. In one case, only one player in the pair could lie, and in the other both players could lie. These are depicted in the columns of Table 1.

As a baseline, piece-rate payoff is implemented: each correct answer gives 1.5 Euros. In the piece-rate, the lies of one party do not harm the other party, and the benefit of each lie is constant. We further varied whether only one person can lie or both parties can lie in the piece-rate. The payoff structures of the two treatments are depicted in the second row of Table 1.

---

[7] We increased this number in the later experimental sessions in accordance with the increasing number of observations we had for the correct answers. The rationale for such a change was to correct for the small sample size errors. All results reported in this paper are standardized and take ratios into account.

[8] We were not explicit in our wording regarding the no-lie condition. Our exact wording was as follows: "From an earlier session of this experiment, we know the number of correct answers made by all the subjects." We cannot rule out that some subjects considered declared values rather than the true values when answering. None of the subjects asked for a clarification, and they took this guessing part seriously and spent considerable time on it. Furthermore, as we show in section 5, we do not find any effect of treatment in subjects' answers, nor an effect of whether they can lie.

In the pie treatments, players earned a portion of a fixed amount of money according to the ratio of their own declaration to the total declaration within the pair. We chose 15 Euros as the fixed amount based on the results of our pilot which showed approximately five correct answers per person[9]. We then multiplied five answers by the piece-rate payoff of 1.5 Euros for two persons. In the pie treatment, lying increases one's payoffs at a cost to the other, and the benefit from each lie and the cost to the other party is the same. However, the marginal cost of a lie to the other party is decreasing in the number of lies. The pie payoff structure is summarized in the third row of Table 1.

Finally, the third payoff structure is the tournament as commonly implemented in the literature, and is depicted in the fourth row of Table 1. In the tournament, the party with the higher number of correct answers becomes the winner, and earns 3 Euros per correct answer whereas the one with the lower number of correct answers earns nothing. The multiplier of three is chosen to equalize the expected payoff of the persons in the median of the correct answer distribution to that of the piece-rate[10]. Unlike the pie treatment, in the tournament, the benefit of a lie and its cost to the other party is discontinuous. As long as one's declaration is lower than the matched partner's declaration, lying has no benefits nor has any costs. If lying changes the winner, then the cost to the other party is quite high (1.5 Euros times the other's declaration), but it also gives a high benefit of 3 Euros per declaration. Lastly, if the losing party tie-breaks by lying, its cost to the other party is equal to its benefit (1.5 Euros times the declaration).

We know from previous experiments that not all subjects lie, and maximal lying is rare. Thus, a substantial proportion of subjects are considered to exhibit lie-averse preferences. We also know that, in many situations, people act reciprocally. Thus, if our premise of lying as a fairness tool is correct, and if subjects anticipate lying by the other party, then we should observe more lying in the two-party treatments compared to one-party treatments under competition[11]. Further, we conjecture that the effect of reciprocity is highest in the tournament payoff because the incentives

[9] As it will be explained in the results section, the true average turned out to be 3.61 correct answers instead of 5. This makes the stakes with the pie size 15 Euros somewhat higher than that of piece-rate.

[10] This expectation is conditional on the matched partner being honest, and the persons in the median of the distribution correctly believing that they are in the median of the answer distribution. Whenever there is variation in the outcome of a task across different persons, tournament cannot give the same expected payoff as in piece-rate for all person's involved. Since in this study we are interested in the effect of different payoff schemes on lying, such a difference is of no primary concern for us.

[11] This prediction would be obtained with a model that takes into account both reciprocity and lie-aversion.

to lie are highest in tournament compared to the other payoff schemes. Pie-sharing gives lower incentives for reciprocal lying because subjects always earn some money, and the benefit of a unit lie is rather small for higher performers. Therefore we would expect a smaller effect of reciprocity in pie-sharing compared to tournament. Finally, piece-rate has no reciprocity component, rather the only possible effect is via the weakening of a social norm. If we see any effect of controlling who can lie, then this would be a clear demonstration of the importance of social norms when lying.

**Table 1 :** Payment structures in different treatments

| | **One-party (OP)** | **Two-party (TP)** |
|---|---|---|
| **Piece rate** | $P_A = 1.5 * Claim_A$<br>$P_B = 1.5 * Real_B$ | $P_A = 1.5 * Claim_A$<br>$P_B = 1.5 * Claim_B$ |
| **Pie** | $P_A = 15 * \frac{Claim_A}{Claim_A + Real_B}$<br>$P_B = 15 * \frac{Real_B}{Claim_A + Real_B}$ | $P_A = 15 * \frac{Claim_A}{Claim_A + Claim_B}$<br>$P_A = 15 * \frac{Claim_B}{Claim_A + Claim_B}$ |
| **Tournament** | $P_{winner} = 3 * Claim(Real)_{Winner}$<br>$P_{loser} = 0$<br>$P_{tie-break} = 1.5 * Claim$ | $P_{winner} = 3 * Claim_{Winner}$<br>$P_{loser} = 0$<br>$P_{tie-break} = 1.5 * Claim$ |

## 4. Results on Lying

The performance of subjects who cannot lie is depicted in Table 2. With no lying, the average number of correct answers is 3.61. The dispersion of the performance is rather high, with a standard deviation of 2.03. We can see from the frequency distribution that about 85 percent of the correct answers are less than or equal to 5. The median of the distribution is 4 correct answers. The mode of the distribution is five; almost one fifth of the subjects had five correct answers. Notice that there is only one person out of 152 subjects who did nine correct, and no one solved more than nine questions correctly.

Table 3 depicts the average declarations per treatment. The type of competition is listed in rows, and who can lie in the columns. The comparison of the declarations of each treatment with the no-lie condition is stated below the corresponding treatment values. The last column compares the column treatments. All p-values are when using one-sided Mann-Whitney exact test.

**Table 2.** Distribution of correct answers without lying

| Correct Answers | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| Percentage of Subjects | 5.3 | 12.5 | 15.1 | 14.5 | 17.1 | 19.7 | 7.2 | 5.3 | 2.6 | 0.7 |

Average=3.61, Standard Deviation= 2.03, N=152

A first observation is that overall, there is very little lying. Most lying happens in the piece-rate, and even then, the average amount of lies are approximately 1.40 answers. Given that there were 14 questions, the amount of lies are about 14 percent of the total possibility. Furthermore, the declarations in the tournament treatment as well as the pie treatment with only one party lying are not significantly different than no-lie condition. Lying in the pie treatment only happens when both parties can lie, and the average declaration is 1.20 units higher than the No-lie condition. When we look at the effect of reciprocity, we only see a weak effect in pie-sharing. Moreover, in the tournament treatment, the effect is in the opposite direction than expected.

**Table 3–**Declarations per treatment

| | | | |
|---|---|---|---|
| **No-Lie** | 3.61 (2.03)<br>N=152 | | |
| | **One-Party** | **Two-Party** | **p*** |
| **Piece-rate** | 4.90 (2.90)<br>N=41 | 5.11 (3.65)<br>N=38 | 0.470 |
| | *p= 0.010* | *p=0.017* | |
| **Pie** | 4.23 (3.53)<br>N=40 | 4.81 (2.83)<br>N=42 | 0.074 |
| | *p= 0.407* | *p=0.012* | |
| **Tournament** | 4.26 (2.84)<br>N=42 | 3.84 (3.02)<br>N=38 | 0.869 |
| | *p=0.110* | *p=0.355* | |

*Mann-Whitney exact test, one-tailed.

The comparison of the declarations across the payoff schemes show that tournament has the lowest level of lies (statistically indistinguishable from the No-lie condition) whereas piece-rate has the highest. The difference between pie and tournament is not significant. We will provide further support for these results in the next section.

Finally, we can try to estimate the percentage of lying across different treatments. Although our design does not let us know the exact amount of lies, we can infer from the No-lie distribution that any declaration that is 9 or higher is almost surely a lie. This gives a lower bound on the percentage of liars. In Table 4, we report the number of persons who declared a number between 9 and 14 in each treatment and the cumulative percentage of those people. One can see that the highest percentage of lying is in the Pie Two-Party treatment with 14.3 percent, and the lowest is in the one-party tournament treatment with only 1 person out of 42. If we leave out these persons who declared 9 or higher, and repeat the statistical test on the comparison of each treatment with the No-Lie treatment, we see that none of the treatments turn out to be significantly different than the No-Lie treatment[12]. Such a result implies that lying by declaring less than 9 happens at most in few cases, therefore the percentage of liars are close to the cumulative percentages reported in Table 4.

**Table 4.** Frequency of declaring 9 or higher

| | **9** | **10** | **11** | **12** | **13** | **14** | **Cumulative Percent** |
|---|---|---|---|---|---|---|---|
| **No-Lie**<br>N=152 | 1 | 0 | 0 | 0 | 0 | 0 | 0.7 |
| **Piece-rate One-Party**<br>N=41 | 1 | 1 | 1 | 2 | 0 | 0 | 12.2 |
| **Piece-rate Two-Party**<br>N=38 | 0 | 1 | 0 | 1 | 0 | 3 | 13.2 |
| **Pie One-Party**<br>N=40 | 0 | 0 | 0 | 0 | 0 | 3 | 7.5 |
| **Pie Two-Party**<br>N=42 | 2 | 1 | 2 | 1 | 0 | 0 | 14.3 |
| **Tournament One-Party**<br>N=42 | 0 | 0 | 0 | 0 | 0 | 1 | 2.4 |
| **Tournament Two-Party**<br>N=38 | 0 | 1 | 0 | 1 | 0 | 1 | 7.9 |

**Gender Results**

We first look at whether there is a performance difference between the genders. The averages per gender in the No-lie condition are reported in Table 5. Due to a mistake with labelling the

[12] The relevant p-values from the comparison with the No-lie treatment with one-sided Mann-Whitney exact test while excluding declarations 9 or higher are as follows: Piece-Rate One-Party, 0.102; Piece-Rate Two-Party, 0.164; Pie One-Party, 0.332; Pie Two-Party, 0.167; Tournament One-Party, 0.141; Tournament Two-Party, 0.128.

questionnaires in one of our sessions, we have in total 122 observations for gender. We can see that women perform significantly worse than men and the difference is about 0.9 questions.

**Table 5.** Averages per gender in the No-Lie condition

| Average | Females | Males |
|---|---|---|
| 3.61 | 3.12 | 4.01 |
| (2.03) | N=60 | N=62 |
| N=152 | | |
| | Two-tailed Mann-Whitney exact test p=0.005 | |

To investigate whether women and men behave differently under different payment schemes, we run a linear regression to explain the difference between a person's outcome from the average in the No-Lie condition of his or her gender controlling for gender, payment scheme and the interaction effects. This way, we can test whether there is a difference in lying behaviour between men and women at the same time taking into account the difference in their performance. Thus, the regression is as follows:

$$Declaration_i - AvgNoLie_{gender(i)} = \alpha + \beta Gender + \gamma Piece + \delta Pie + \eta Reciprocity + \theta Interaction\ Terms + \varepsilon$$

If people of different genders lie significantly differently, then we expect to see the gender variable to have a significant effect. Moreover, if different genders behave differently under different schemes, we expect to see the interaction effect of the treatment with gender to be significant. Table 6 reports the results of the regression that includes (Model I) and excludes reciprocity (Model II). The first column in the table depicts the variable name, the estimated coefficient and the significance results of each model are below the models. Notice that this specification results in all comparisons being made with the tournament.

The regression including reciprocity terms (Model I) shows no significant effect of any of the variables, not even the intercept that establishes that there is lying. Dropping the reciprocity terms gives us Model II. We can then see that without the reciprocity terms, lying becomes weakly significant (intercept p-value=0.072), and there is no effect of the type of competition nor gender. The only weak effect comes from the fact that women lie more in the piece-rate than men (p=0.082).

Finally, to clarify how lying depends on gender, we include the averages with respect to men and women using the pooled data for each payoff scheme in Table 7. We also report the relevant two-sided Mann-Whitney exact test results. We can see that while men are unresponsive to the payoff scheme and lie about the same rate in all treatments, women only lie in the piece-rate.

**Table 6.** Regression Results

| | Model I | | | Model II | | |
|---|---|---|---|---|---|---|
| | Coefficient | Std. Error | Sig | Coefficient | Std. Error | Sig |
| Intercept | 0.935 | 0.762 | .221 | 0.935 | 0.517 | .072 |
| Female | -1.052 | 1.001 | .295 | -0.608 | 0.685 | .376 |
| Piece-Rate | -0.333 | 1.459 | .819 | -0.258 | 0.748 | .731 |
| Pie | 0.077 | 0.968 | .937 | 0.091 | 0.688 | .895 |
| Reciprocity | 0.000 | 1.047 | 1.000 | - | | |
| Female x Piece-Rate | 2.333 | 2.024 | .250 | 1.909 | 1.093 | .082 |
| Female x Pie | 1.298 | 1.393 | .352 | 0.384 | 0.960 | .690 |
| Female x Reciprocity | 0.870 | 1.386 | .531 | - | | |
| Piece-Rate x Reciprocity | 0.093 | 1.736 | .957 | - | | |
| Pie x Reciprocity | 0.034 | 1.403 | .981 | - | | |
| Female x Piece-Rate x Reciprocity | -0.830 | 2.451 | .735 | - | | |
| Female x Pie x Reciprocity | -1.707 | 1.954 | .383 | - | | |
| | $R^2$=0.029 | | | $R^2$=0.022 | | |

**Table 7.** Gender averages per payoff scheme

| | Female | Male | p* |
|---|---|---|---|
| **Piece-rate** | 5.10 (3.36)<br>N=21 | 4.74 (2.90)<br>N=31 | 0.805 |
| **Pie** | 3.92 (2.92)<br>N=37 | 5.09 (3.34)<br>N=44 | 0.089 |
| **Tournament** | 3.44 (2.37)<br>N=45 | 5.00 (3.30)<br>N=34 | 0.040 |

* Two-sided Mann-Whitney exact test

## 5. Beliefs:

We elicited two types of beliefs after the experiment: beliefs about fairness and beliefs about the distribution of correct answers. To elicit the subjects' fairness ideas, we asked how a fixed amount of money (15 Euros) should be distributed within a pair assuming different combinations of correct answers. In total they were asked to state 12 choices with the following correct answers within a pair: (14,0), (12,2), (10,4), (8,6), (7,7), (14,7), (6,5), (6,4), (6,3), (6,2), (6,1), (6,0). Since there is almost no variation in the subjects' choices for (7,7)[13], it is dropped from our subsequent analysis. To elicit the subjects' beliefs about the distribution of correct answers, we asked them to guess the average number of correct answers. Additionally, we asked them to estimate how many subjects answered 0 question correctly, 1 question correctly,…, 14 questions correctly. We incentivized the answers by paying 6 Euros for a fully correct estimation with 50 eurocents reduction per deviation. If there were 12 or more differences, the earnings were zero. We standardized their answers to ratios. By looking at the subjects' answers, we can also calculate the average of the distribution they guessed. We also report the difference between the estimated distribution average and their guessed average. The mean and the standard deviation of all the variables are included in the Table A1 in the Appendix.

We exclude the fair division (7,7) and the difference between the averages from further analysis, and are therefore left with 28 variables. A first observation is that none of these variables are differently distributed across treatments (using two-sided Mann-Whitney exact tests)[14]. Therefore, we can use the data from all the treatments. Secondly, the 28 variables are highly correlated with one another in a specific pattern: As expected, 'fair division' answers and the distributional guesses are not correlated. The fair division answers highly correlate with each other. Also, the guesses on the distribution of correct answers highly correlate with each other and with the guessed average. This leaves room for factor analysis so that we can reduce the number of variables in a way that explains the most variance. The rotated matrix of the factor analysis as well as variance explained is reported in the Appendix. The resulting number of factors are six; four variables mostly consist of the estimation of the distribution of correct answers and the averages, and the other two variables are about the division of money.

Using these factors, we study the relationship between declarations and estimated ability distribution as well as the relationship between ability and fairness considerations. We know from the Dunning–Kruger effect (Dunning and Kruger, 1999; Schlösser, Dunning, Johnsonand Kruger, 2013) that people tend to think others are like them when judging the ability distribution. Thus,

[13] The reason that we included (7,7) was to capture concave preferences that value an extreme distribution over an equal one. We did not find any evidence for such preferences.

[14] Given that there are 28 variables per treatment, there are 168 comparisons in total, therefore the threshold for significance have to be adjusted by 1/168 (Bonferroni adjustment). This is necessary to rule out finding significance due to a large number of tests. With the usual p-level of 0.05, 11 of these comparisons show up as significant, but none of them survive the adjusted threshold.

low ability people underestimate the ability of others whereas high ability people overestimate the ability of others. If such an effect exists in our experiment, then we would expect to see a reflection of it in the related factors.

The second relationship that we can investigate is the one between perceptions of fairness and ability. There is some research that study whether fairness considerations take effort into account (see for example Almås, Cappelen, Sørensen, and Tungodden, 2010; Cappelen, Hole, Sørensen, and Tungodden, 2007) however to our knowledge there is no study that investigates whether ability and fairness considerations are correlated. We expect higher ability people to think that higher ability people should get a large share of the pie whereas low ability people to opt for a fairer share of the pie. Again, such an effect will be captured in the relevant factors.

**Table 8.** Regression results using factors

| | Coefficient | Std. Error | Significance |
|---|---|---|---|
| Intercept | 0.923 | 0.525 | 0.081 |
| Factor 1 | 1.073 | 0.197 | 0.000 |
| Factor 2 | 0.431 | 0.205 | 0.037 |
| Factor 3 | 0.261 | 0.192 | 0.177 |
| Factor 4 | -0.004 | 0.178 | 0.981 |
| Factor 5 | 0.452 | 0.212 | 0.035 |
| Factor 6 | -0.164 | 0.190 | 0.387 |
| Female | -0.574 | 0.676 | 0.397 |
| Piece-Rate | -0.753 | 0.738 | 0.309 |
| Pie | 0.195 | 0.684 | 0.776 |
| Female*Piece-Rate | 2.017 | 1.048 | 0.056 |
| Female*Pie | 0.711 | 0.930 | 0.446 |

$R^2$=0.187

Table 8 depicts the regression results using the variables derived from factor analysis. The dependent variable is, as in the previous regressions, the difference between one's declared performance and the mean of his or her gender in the No-Lie condition. The control variables are the six factors, gender, treatment, and gender and treatment interaction effect. Because the magnitude of the factored variables are not easy to interpret, we will only focus on the sign of the estimated coefficients.

We can see from the table that the inclusion of the factors explain a substantial amount of variance in the independent variable, and the $R^2$ increases from 0.022 to 0.187. Among the six factors, the first, second and the fifth have a significantly positive effect on the difference of the declarations from the baseline.

Factor 1 is negatively correlated with the estimated ratio of the persons in the population with 0, 1, 2, 3, 4 correct answers, and positively correlated with the ratio of the persons with 6, 7, ..11 correct answers. The significantly positive coefficient of the first factor in the regression suggests that people who declare higher than the *real average* of their gender (3.1 for females, 4.0 for males) guess a distribution to the right of the *average estimated* distribution (centred at 5 correct answers), and persons who declare lower than the *real average* of their gender guess a distribution to the left of the *average estimated* distribution. To put it differently, in any payoff scheme, any person who thinks most others did four or fewer correct answers declares less than the average true ability of their gender, and any person who thinks that most others did six or more declares more. This effect might be a result of two distinct effects occurring at the same time: overestimation of the population ability, and judging others' performance to be close to one's own performance (Dunning-Kruger effect). Notice that, just overestimating the population ability is not enough to get the positive effect of Factor 1: Assume everyone estimates the real average to be five; if own performance was not correlated with one's estimation of the population performance, then Factor 1 would not have any explanatory power on declared performances. However, the regression in Table 8 alone cannot tell us whether the latter effect is related to lying or only stems from the misjudgement of others' ability. We will further analyze this in the next subsection.

Factor 5 also has a significantly positive effect. This factor is related to the guesses of 0 and 1 correct answers (negative correlation) and 4 and 5 correct answers (positive correlation) pointing to a different type of participant than the one captured by Factor 1. This second type guesses about three correct answers, and they are closer to the real distribution of the population. Since this factor has a significantly positive effect, this type also estimates others' performance to be close to own performance.

Answers to the fair distribution questions also help explain declarations. The effect of the second factor is significantly positive, and this factor is mainly driven by the answers to how to distribute

15 Euros when a pair of players have (14,0), (12,2), (10,4), (6,4), (6,3), (6,2), (6,1), (6,0) correct answers. The positive effect provides evidence that people with higher declarations think people with higher declarations should earn more. We will further investigate whether there is any systematic difference in the fairness factor with or without lying.

Finally, we can confirm from this regression that there is no systematic significant effect of gender in the amount of lies nor the type of competition except in the piece-rate.

**Relationship between beliefs, fairness and lying:**

If lying is not dependent on ability, then we would expect everyone to lie about the same rate, and the two effects observed in the regression of Table 8 would also be observed when the observations come from the No-Lie subjects. If, on the other hand, low and high ability people lie at different rates, we would see a difference in the effect of factors with No-lie subjects. Similarly, if there is a correlation between what one considers fair and the rate of lying, we would not be able to observe that effect by only looking at lying behaviour. To test whether lying is dependent on ability or is correlated with fairness attitudes, we have to compare the estimated effects of our factors in explaining the real performance in the No-lie condition to the declared performance in the lie treatments. The No-Lie treatment gives us the population estimates. Any difference in the estimated coefficients between the No-Lie treatment and the lie treatments standardized with the estimated standard errors is approximately distributed with a t-distribution[15].

From Table 9, we can see that most of the estimated coefficients in both the No-Lie and lie treatments are similar. Among the six factors, only the third factor is significantly different in the two conditions with a two-sided t-test p=0.039. Since this third factor is driven by the estimated ratios of persons who declare 9 or higher, there is evidence that people who estimate a relatively high percentage of high declarations declare high values themselves only in the lying treatments. However, this third factor has no significant effect in either of the regressions. Furthermore, none of the factors that significantly contributed to explaining the variance in declarations have a significantly different coefficient in the No-Lie and lie treatments. Therefore, we can conclude that there is no indication that lying depends on ability nor on what is considered fair.

[15] Formally, if the estimated coefficient is $\hat{\beta}$ and the standard error is $\hat{\sigma}$ from the Lie treatments, and the estimated coefficient is β from the No-Lie treatment, then $\widehat{(\beta - \beta)}/\hat{\sigma}$ is distributed with a t-distribution with (Number of observations-Number of variables-1) degrees of freedom.

**Table 9.** Separate Linear Regressions

| | No-Lie Condition | | | Lie Treatments | | |
|---|---|---|---|---|---|---|
| | Coefficient | Std. Error | Significance | Coefficient | Std. Error | Significance |
| Constant | 3.235 | 0.202 | 0.000 | 4.054 | 0.281 | 0.000 |
| Factor 1 | 0.985 | 0.143 | 0.000 | 1.057 | 0.196 | 0.000 |
| Factor 2 | 0.322 | 0.139 | 0.022 | 0.459 | 0.205 | 0.026 |
| Factor 3 | -0.189 | 0.158 | 0.235 | 0.206 | 0.190 | 0.280 |
| Factor 4 | 0.054 | 0.192 | 0.780 | -0.035 | 0.177 | 0.841 |
| Factor 5 | 0.541 | 0.131 | 0.000 | 0.426 | 0.211 | 0.044 |
| Factor 6 | -0.365 | 0.160 | 0.024 | -0.156 | 0.190 | 0.414 |
| Male | 0.882 | 0.287 | 0.003 | 0.800 | 0.403 | 0.049 |
| | N=115, $R^2$=0.465 | | | N=200, $R^2$=0.188 | | |

## 6. Discussion and Conclusion

In this paper, we experimentally investigate whether lying is used as a tool for fairness. We introduced a real effort (analytical) task that gave some people the opportunity to lie. To find out whether cheating takes place to restore equity, we compared lying in the one-party treatments in which only one person in the pair can lie and to the two-party treatments in which both subjects in the pair can lie.

We found that overall, only few people lie, and most lying is at an intermediate level. Our results showed that reciprocity plays no role in lying behaviour. Furthermore, somewhat at odds with the results of previous studies, we found that most lying occurs in the piece-rate and that there is no evidence for lying in the tournament. Finally, in our analytical task, women perform worse than men, but they lie more only in the piece-rate. There is no evidence for women lying in the pie and tournament payoffs. Men, however, seem unresponsive to the payment scheme.

Further analysis on the subjects' estimations of the performance distribution showed a systematic effect of own ability on one's estimation of others' ability. The results are in line with the Dunning–Kruger effect; lower performers underestimate others' performance and higher performers overestimate. Finally, we found evidence for the fairness judgments to be negatively correlated with ability. That is, higher performers propose more unequal distributions as an outcome of the task.

The discrepancy between the lying rates in our setup and the ones reported in the literature might stem from our task choice. It is likely that expectations on others’ cheating behaviour depend on the task. A direct test of how beliefs about others’ lies affect one’s own lying behaviour in different tasks is left for future research. Other aspects of the task might also matter. It is not inconceivable that a coin flipping or die-throwing task makes it easier to lie not only because such tasks do not capture any ability but also because subjects perceive it unfair to be paid by the outcome of a random device. Tasks that reveal information about one’s ability in a domain that is generally valued might trigger honesty. Lying to the experimenter might also be perceived as different than lying to another subject. Finally, the setup of the experiment might matter: Whether cheating possibility is too obvious or explicitly stated in the experiment can have a large influence on the outcomes, and should be taken into consideration in future research.

**Acknowledgements**

We gratefully acknowledge the financial support by the Research Priority Area of Behavioral Economics of the University of Amsterdam. We thank Rudy Ligtvoet for statistical support. This article has benefitted from the comments by Theo Offerman, and Jeroen van de Ven.

**References**


Almås, I., Cappelen, A. W., Sørensen, E. Ø., and Tungodden, B. (2010). Fairness and the development of inequality acceptance. *Science*, *328*(5982), 1176-1178.

Ariely, Kamenica, and Prelec, (2008). “Man's search for meaning: the case of Legos”, Journal of Economic Behavior and Organization, 67 (2008), pp. 671–677

Banchero, (28 January 2014). Columbus, Ohio, School District Hit By Cheating Allegations. The Wall Street Journal. Retrieved from http://online.wsj.com/news/articles/SB10001424052702303277704579349202004452142

Becker, W., H.J. Buchner, S. Sleeking (1987). “The impact of public transfer expenditure on tax evasion: An experimental approach” Journal of Public Economics, 34, pp. 341–350.

Belot, M., & Schröder, M. (2013). Sloppy work, lies and theft: A novel experimental design to study counterproductive behaviour. *Journal of Economic Behavior & Organization*, *93*, 233-238.

Cappelen, Alexander W., Astri Drange Hole, Erik Ø Sørensen, and Bertil Tungodden. (2007). "The Pluralism of Fairness Ideals: An Experimental Approach." *American Economic Review*, 97(3): 818-827.

Carpenter, J., P. Matthews, and J. Schirm (2010), "Tournaments and Office Politics: Evidence from a real effort experiment,” *American Economic Review*, 100(1), 504-517.

Charness, G., Masclet, D., & Villeval, M. C. (2013). The dark side of competition for status. M*anagement Science*, *60*(1), 38-55.

Gill, David, Victoria Prowse, and Michael Vlassopoulos (2013). “Cheating in the workplace: An experimental study of the impact of bonuses and productivity”, Journal of Economic Behavior & Organization, Volume 96, December 2013, Pages 120-134,

Dato, S., and Nieken, P. (2014). Gender differences in competition and sabotage. *Journal of Economic Behavior & Organization*, *100*, 64-80.

Erat, S., and Gneezy, U. (2012). White lies. *Management Science*, *58*(4), 723-733.

Falk, A., and Szech, N. (2013). Morals and markets. *Science*, 340(6133), 707-711.

Fischbacher, U., and Heusi, F., (2008). Lies in disguise: an experimental study on cheating. Thurgau Institute of Economics Research Paper No. 40.

Gibson, Rajna, Carmen Tanner, and Alexander F. Wagner. (2013), "Preferences for Truthfulness: Heterogeneity among and within Individuals." *American Economic Review*, 103(1): 532-48.

Gneezy, U., 2005. Deception: the role of consequences. American Economic Review 95 (1), 384–394.

Gneezy, U., Rockenbach, B., & Serra-Garcia, M. (2013). Measuring lying aversion. *Journal of Economic Behavior & Organization*, *93*, 293-300.

Hurkens, S., and N. Kartik. (2009). Would I lie to you? On social preferences and lying aversion. *Experimental Economics*, *12*(2), 180-192.

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* 77.6: 1121.

Lundquist, T., Ellingsen, T., Gribbe, E., and Johannesson, M. (2009). The aversion to lying. *Journal of Economic Behavior & Organization*, *70*(1), 81-92.

Matsaganis, M., & Flevotomou, M. (2010). Distributional implications of tax evasion in Greece. GreeSE Paper No. 31

Moore, M. (20 June 2013). Riot after Chinese teachers try to stop pupils cheating. *The Telegraph*. Retrieved from: http://www.telegraph.co.uk/news/worldnews/asia/china/10132391/Riot-after-Chinese-teachers-try-to-stop-pupils-cheating.html

Oxoby, R. J., & Spraggon, J. (2008). Mine and yours: Property rights in dictator games. *Journal of Economic Behavior & Organization*, *65*(3), 703-713.

Pascual-Ezama, D., Prelec, D., & Dunfield, D. (2013). Motivation, money, prestige and cheats. *Journal of Economic Behavior & Organization*, *93*, 367-373.

Rigdon, Mary L. and D'Esterre, Alexande (2012). The Effects of Competition on the Nature of Cheating Behavior. Retrieved from SSRN: http://ssrn.com/abstract=2194870.

Schlösser, T., Dunning, D., Johnson, K. L., and Kruger, J. (2013). How unaware are the unskilled? Empirical tests of the “signal extraction” counterexplanation for the Dunning–Kruger effect in self-evaluation of performance. *Journal of Economic Psychology*, *39*, 85-100.

Schwieren, C., and Weichselbaumer, D. (2010). Does competition enhance performance or cheating? A laboratory experiment. *Journal of Economic Psychology*, *31*(3), 241-253.

Telegraph Sport (2013, January 18) Lance Armstrong's interview with Oprah Winfrey: the transcript. The Telegraph. Retrieved from http://www.telegraph.co.uk/

Wallschutzky, I. G. (1984). Possible causes of tax evasion. *Journal of economic psychology*, *5*(4), 371-384.

Wenzel, M. (2004). The social side of sanctions: personal and social norms as moderators of deterrence. *Law and Human Behavior*, *28*(5), 547.

**Appendix. Table A1.** Averages of fairness answers and estimated distribution

| Fairness Answers | N | Mean (Std Dev) |
|---|---|---|
| Division 14,0 | 338 | 13.78 (2.05) |
| Division 12,2 | 337 | 12.01 (1.74) |
| Division 10,4 | 337 | 10.38 (1.46) |
| Division 8,6 | 337 | 8.68 (1.25) |
| Division 7,7 | 337 | 7.50 (0.45) |
| Division 14,7 | 335 | 10.06 (1.46) |
| Division 6,5 | 336 | 8.44 (3.89) |
| Division 6,4 | 336 | 8.89 (1.29) |
| Division 6,3 | 336 | 9.71 (1.38) |
| Division 6,2 | 336 | 10.63 (1.63) |
| Division 6,1 | 335 | 11.78 (1.87) |
| Division 6,0 | 336 | 13.15 (2.53) |

| Estimated Distribution | N | Mean (Std Dev) |
|---|---|---|
| Average Correct Answers | 347 | 5.02 (1.69) |
| Percentage of 0 correct | 346 | 0.04 (0.07) |
| Percentage of 1 correct | 346 | 0.06 (0.08) |
| Percentage of 2 correct | 346 | 0.10 (0.08) |
| Percentage of 3 correct | 345 | 0.13 (0.09) |
| Percentage of 4 correct | 345 | 0.15 (0.08) |
| Percentage of 5 correct | 346 | 0.14 (0.09) |
| Percentage of 6 correct | 346 | 0.12 (0.08) |
| Percentage of 7 correct | 345 | 0.09 (0.07) |
| Percentage of 8 correct | 346 | 0.06 (0.05) |
| Percentage of 9 correct | 345 | 0.04 (0.04) |
| Percentage of 10 correct | 346 | 0.03 (0.03) |
| Percentage of 11 correct | 346 | 0.02 (0.02) |
| Percentage of 12 correct | 346 | 0.01 (0.02) |
| Percentage of 13 correct | 345 | 0.01 (0.02) |
| Percentage of 14 correct | 346 | 0.01 (0.02) |
| Average from Estimation | 341 | 4.96 (1.53) |
| Difference between Avg’s | 341 | -0.07 (1.21) |

**Appendix-Factor Analysis**

**Rotated Component Matrix**

| | Component | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Division140 | .062 | .849 | -.090 | .022 | .106 | .100 |
| Division122 | .035 | .826 | -.103 | .229 | .172 | .107 |
| Division104 | .004 | .648 | -.050 | .482 | .209 | .142 |
| Division86 | -.053 | .297 | .011 | .772 | .154 | .147 |
| Division147 | .034 | .249 | -.010 | .572 | .168 | .221 |
| Division65 | .072 | -.016 | -.059 | .493 | -.102 | -.243 |
| Division64 | .028 | .397 | .051 | .791 | -.096 | -.075 |
| Division63 | .066 | .586 | .039 | .647 | -.133 | -.117 |
| Division62 | .049 | .720 | -.002 | .427 | -.085 | -.137 |
| Division61 | .025 | .835 | .034 | .234 | -.063 | -.056 |
| Division60 | -.047 | .883 | -.004 | .078 | -.075 | -.090 |
| Average | .680 | .009 | .300 | -.017 | .215 | -.072 |
| Updated0 | -.334 | -.082 | -.036 | -.005 | -.732 | .060 |
| Updated1 | -.500 | .002 | -.141 | -.048 | -.632 | .223 |
| Updated2 | -.715 | -.052 | -.206 | -.035 | -.216 | .400 |
| Updated3 | -.773 | .031 | -.237 | -.027 | .237 | .248 |
| Updated4 | -.527 | -.010 | -.238 | .002 | .576 | -.251 |
| Updated5 | .136 | .050 | -.215 | .002 | .326 | -.767 |
| Updated6 | .717 | -.077 | -.174 | .079 | .164 | -.256 |
| Updated7 | .870 | .036 | -.055 | .056 | .075 | .013 |
| Updated8 | .812 | .085 | .159 | .030 | .033 | .233 |
| Updated9 | .708 | .057 | .363 | -.038 | .025 | .246 |
| Updated10 | .473 | .091 | .667 | -.048 | .007 | .227 |
| Updated11 | .331 | .108 | .765 | .030 | -.012 | .188 |
| Updated12 | .133 | .011 | .907 | .063 | -.015 | .039 |
| Updated13 | .118 | -.105 | .861 | -.053 | -.019 | -.007 |
| Updated14 | .018 | -.187 | .712 | -.027 | .020 | -.059 |
| UpdatedAvg | .791 | .034 | .523 | .027 | .293 | -.058 |

**Total Variance Explained**

| Component | Rotation Sums of Squared Loadings | | |
|---|---|---|---|
| | Total | % of Variance | Cumulative % |
| 1 | 5.676 | 20.272 | 20.272 |
| 2 | 4.568 | 16.313 | 36.585 |
| 3 | 3.908 | 13.957 | 50.542 |
| 4 | 2.768 | 9.884 | 60.427 |
| 5 | 1.834 | 6.550 | 66.977 |
| 6 | 1.427 | 5.098 | 72.075 |

**Appendix. (Instructions for Tournament Two-Party)**

**General instructions**

Welcome to this experiment. We will first go through the instructions together. Talking is strictly forbidden during the experiment, if you have any questions please raise your hand and we will come to your table to answer your question.

The experiment will last about 45 minutes. The experiment will consist of two parts. Below you find the instructions for the first part. After the first part is completed, you will receive instructions for the second part. At the end of the experiment, we ask you to fill in a short questionnaire. Your payment in this experiment is based on the task that you do in the first part. The determination of your payment is conducted in the second part.

In this experiment you will be randomly given the label of player A, player B or player C. Your player label is indicated on your seat. At the beginning of the experiment, every player A (B) will be randomly matched with another player B (A). You will not learn with whom you are matched with during or after the experiment. Your player labels will stay the same throughout the whole experiment. Additionally, the person with whom you are matched will stay the same.

**First part**

<u>Instructions for player As and player Bs:</u>

In this part, all player As and all player Bs will be given a document with questions, and an answer sheet. The document contains the same questions for everyone. Your answers to these questions form the basis of your payment; thus, you should find the correct answers to as many questions as possible and write them down on the answer sheet. Each question has only one correct answer. You have 10 minutes to indicate your answers. In this part, the exact order of events for player As and player Bs is as follows:

→ You will receive two documents with blank pages facing up. You should keep these documents as they are until the experimenter tells you to start.
→ When the experimenter says “You can start”, you can turn the documents around, and start solving the questions. You then have 10 minutes to solve as many questions as possible. Note down your answers in the answer sheet. Notice that only the answers noted down on the answer sheet matters.
→ When the 10 minutes are over, the experimenter will say “Time is up”. You then have to put down your pen and turn the documents around so that blank pages are facing up.
→ An experimenter will collect all the documents without turning them and hand them to the randomly assigned player Cs. Player Cs will not know which document belongs to which seat number.
→ Player Cs will check the answers on the answer sheets, and denote the correct answers with highlighter.
→ Player Cs will hand in the checked answer sheets blank pages facing up to the experimenter.
→ An experimenter will hand the answer sheets back to player As and player Bs.
→ You will receive the instructions for part 2.

Instructions for player Cs:

In this part, all player As and all player Bs will be given a document with questions. This document contains the same questions for everyone. Each question has only one correct answer. Player As and player Bs will have 10 minutes to indicate their answers. Their answers will be collected by an experimenter and some of them will be handed in to you. You will not know which document belongs to which seat number. In this part, the exact order of events relevant for player Cs is as follows:

→ Player As and player Bs will receive their documents and will be given 10 minutes to solve as many questions as possible.
→ You will be given a sheet indicating the correct answers to each question.
→ When 10 minutes is over, an experimenter will hand in to you some of the documents with blank pages facing up. You will not know which document belongs to which seat number.
→ You will turn around the document on the top of the pile, check the answers, and only highlight the correct answers with a highlighter.
→ You will turn the document around again so that the blank pages are facing up. You will place the document in the top right corner of your table.
→ You will repeat steps 4 and 5 until there are no more documents left. Beware of the fact that the order of checking answers should be strictly adhered to for the successful execution of this part.
→ An experimenter will collect all the documents.
→ An experimenter will determine your payment from this part as explained below.
→ An experimenter will hand the documents back to player As and player Bs.
→ You will receive the instructions for part 2.

Payment for Player Cs: You will be paid according to the accuracy of checking the answers. For each player C, an experimenter will randomly draw one of the checked documents, and check the corrections made by that player. You will receive 10 Euros if the document is completely correctly checked, and 0 Euros otherwise. This payment will be in addition to the 3 Euro show-up fee.

We will now continue with the second part of this experiment.

**Second Part**

Instructions for player As and player Bs:

In this part, your payment is determined. Your earnings in this experiment will depend upon your individual decisions and the decisions of the player you are matched with.

Your earnings in this experiment are determined as follows: among every matched player A and player B, the player with the higher number of correct answers will earn 3.00 Euros for every correct answer. The player with the lower number of correct answers will earn nothing. If the number of correct answers are the same, then both players will earn 1.50 Euros for every correct answer. You will in addition earn 3 Euros show-up fee:

There are three situations that can arise.

Situation 1:
In situation 1 the $\text{Number of Correct Answers}_{\text{A}} > \text{Number of Correct Answers}_{\text{B}}$, then the total earnings of player A and player B is as follows:

Player A: $\textbf{Total Earnings}_{\textbf{A}}\textbf{=}\ \text{Number of Correct Answers}_{\text{A}} * \mathbf{3+3}$
Player B: $\textbf{Total Earnings}_{\textbf{B}}\textbf{=}\ \mathbf{3}$

Situation 2:
In situation 2 the $\text{Number of Correct Answers}_{\text{B}} > \text{Number of Correct Answers}_{\text{A}}$, then the total earnings of player A and player B is as follows:

Player A: $\textbf{Total Earnings}_{\textbf{A}}\textbf{=}\ \mathbf{3}$
Player B: $\textbf{Total Earnings}_{\textbf{B}}\textbf{=}\ \text{Number of Correct Answers}_{\text{B}} * \mathbf{3+3}$

Situation 3:
In situation 3 the $\text{Number of Correct Answers}_{\text{A}} = \text{Number of Correct Answers}_{\text{B}}$, then the total earnings of player A and player B is as follows:

Player A: $\textbf{Total Earnings}_{\textbf{A}}\textbf{=}\frac{\text{Number of Correct Answers}_{\text{A}}}{\mathbf{2}} * \mathbf{3+3}$
Player B: $\textbf{Total Earnings}_{\textbf{B}}\textbf{=}\frac{\text{Number of Correct Answers}_{\text{B}}}{\mathbf{2}} * \mathbf{3+3}$

You can now examine how many correct answers you have. In the meantime, all player As and player Bs will receive a declaration form. All player As and player Bs are required to declare the amount of correct answers on the declaration form, in order for the experimenters to calculate their payment. When you are finished with examining your answer sheet, please fold your declaration forms in two- blank page facing up. Once everyone has checked their corrected answer sheets, an experimenter will collect the declaration forms of all player As and player Bs, and hand it to another experimenter (who is not involved in the running of the experiment) to calculate the payoffs. The answer sheets of all players will be handed back to Player Cs who will then highlight all the answers on the answer sheet. The declaration forms and the answer sheets will then be returned to the players. <u>You can keep</u> the declaration forms and the answer sheet with you after the experiment.

After determining everyone's payoffs, an experimenter will hand out questionnaires. Please fill in the questionnaire as accurately as possible. After everyone finishes the questionnaire, you will be handed in your payment in an envelope with a receipt. <u>Please sign the receipt</u>, and put it back in the envelope. We will then call you one by one; you will proceed to the waiting room where you will hand back your envelope with the signed receipt in, and afterwards you can leave. Please stay seated and remain quiet until the experimenter calls you.

<u>Instructions for player Cs:</u>

In this part we will determine the payoffs of player As and player Bs. You will be handed the answer sheets of player As and player Bs which you will highlight all the answers. The order, as before, is very important. An experimenter will then pick the answer sheets and hand them back to player As and player Bs.

After determining everyone's payoffs, an experimenter will hand out questionnaires. Please fill in the questionnaire as accurately as possible. After everyone finishes the questionnaire, you will be handed in your payment in an envelope with a receipt. <u>Please sign the receipt</u>, and put it back in the envelope. We will then call you one by one; you will proceed to the waiting room where you will hand back your envelope with the signed receipt in, and afterwards you can leave. Please stay seated and remain quiet until the experimenter calls you.

Appendix. (Task)

**1.** A certain jar contains 100 jelly beans – 44 white, 36 green, 11 yellow, 5 red, 4 purple. If a jelly bean is chosen at random, what is the chance that it is neither red nor purple? Give your answer as a minimal fraction, i.e., such as $\frac{a}{b}$ in which a and b are natural numbers without a common divisor larger than 1.

**2.** If 2 workers can complete painting 2 walls in exactly 2 hours, how many workers would be needed to paints 18 walls in 6 hours ?

**3.** A car got 17 kilometers per liter using gasoline that costs €2,00 per liter. What was the cost, in euros, of the gasoline used in driving the car 476 kilometers? Give your answer as a number with two decimals.

**4.** If y=3x and z=2y, what is x+y+z in terms of x?

**5.** A certain shipping company charges an insurance fee of:

€0.75 when shipping any package with contents worth €25.00 or less,
€1.00 when the content is worth more than €25.00, but worth €50.00 or less and
€1.50 when the content is worth more than €50.00.

What is the total insurance fee that Don has to pay when he sends four different packages worth €18.50, €25.00, €27.50, and €57.00?

**6.** In year Y, the population of Colorado was approximately half that of New Jersey, and the land area of Colorado was approximately 14 times that of New Jersey. The population density (number of persons per unit of land area) of Colorado in year Y was approximately how many times the population density of new Jersey? Give your answer as a minimal fraction, i.e., such as $\frac{a}{b}$ in which a and b are natural numbers without a common divisor larger than 1.

**7.** Machine R, working alone at a constant rate, produces x units of a product in 30 minutes, and machine S, working alone at a constant rate, produces x units of the product in 48 minutes, where x is a positive integer.

Quantity A: the number of units of the product that machine R, working alone at its constant rate, produces in 3 hours
Quantity B: the number of units of the product that machine S, working alone at its constant rate, produces in 4 hours.

Given this information, find the quantity A in terms of quantity B?

**8.** Of the people in a certain survey, 58 percent were at most 40 years old and 70 percent were at most 60 years old. If 252 of the people in the survey were more than 40 years old and at most 60 years old, what was the total number of people in the survey?

**9.** Among the 9.000 people attending a football game at college C, there were x students from college C and y students who were not from college C.

<u>Quantity A:</u> The number of people attending the game who were not students
<u>Quantity B:</u> 9.000-x-y

Given this information, what can you decide about the relation between the quantity A and quantity B?

**10.** A manager is forming a 6-person team to work on a certain project. From the 11 candidates available for the team, the manager has already chosen 3 to be on the team. In selection of the other 3 team members, how many possible combinations of the three remaining candidates does the manager have to choose from?

**11.** What is the smaller angle formed by the hands of the clock when it is 1:45?

**12.** Eight points are equally spaced on a circle. If 4 of the 8 points are to be chosen at random, what is the probability that a quadrilateral having the 4 points chosen as vertices will be a square?

**13.** You have a large cube made up of 10 x 10 x 10 smaller cubes. Each of the smaller cubes is identical. You submerge the larger cube entirely in red paint. How many of the smaller cubes have paint on them?

**14.** How many times a day do the hands of a clock make right angles with each other?

## QUESTIONNAIRE

1. Gender male / female

2. Study ……………………………..

3. Suppose player A and B participated in the math sum assignment as in this experiment. In the first two columns, you find the correct answers of A and B respectively in different combinations. Please indicate how 15 Euros of payment should be divided between the two players according to their answers:

| **Results task** | | **Your proposed division** | |
|---|---|---|---|
| 14 | 0 | | |
| 12 | 2 | | |
| 10 | 4 | | |
| 8 | 6 | | |
| 7 | 7 | | |
| 14 | 7 | | |
| 6 | 5 | | |
| 6 | 4 | | |
| 6 | 3 | | |
| 6 | 2 | | |
| 6 | 1 | | |
| 6 | 0 | | |

4. How many questions do you think participants have on average correct?

…………..

For question 5, you earn an additional payment.

5. Next, we ask you to estimate how many questions the other people in the experiment have correct. From earlier sessions of this experiment, we know the number of correct answers made by the subjects. There were in total 100 subjects. Below, we ask you to estimate the number of subjects who answered 0 correct, 1 correct,….,14 correct. You will be given 6 Euros for a fully correct answer. For every difference between your answer and the correct answer, we will subtract 0.5 Euros. If your answer differs by more than 12 from the correct answers your payment will be 0. Make sure that the total number of subjects is 100 !

Number of subjects who had 0 correct answers:
Number of subjects who had 1 correct answers:
Number of subjects who had 2 correct answers:
Number of subjects who had 3 correct answers:
Number of subjects who had 4 correct answers:
Number of subjects who had 5 correct answers:
Number of subjects who had 6 correct answers:
Number of subjects who had 7 correct answers:
Number of subjects who had 8 correct answers:
Number of subjects who had 9 correct answers:
Number of subjects who had 10 correct answers:
Number of subjects who had 11 correct answers:
Number of subjects who had 12 correct answers:
Number of subjects who had 13 correct answers:
Number of subjects who had 14 correct answers:

6. How many experiments have you participated in since September 2012 (previous and this academic year)? …………

7. How many times have you done a GRE or GMAT preparation test at home or in a course?

○ 0-1 times ○ 2-3 times ○ 4-5 times ○ 6 or more