



UNIVERSITÀ
DEGLI STUDI
FIRENZE

FLORE

Repository istituzionale dell'Università degli Studi di Firenze

Experimenter bias and gender difference

Questa è la Versione finale referata (Post print/Accepted manuscript) della seguente pubblicazione:

Original Citation:

Experimenter bias and gender difference / A. Innocenti; M.G. Paziienza. - STAMPA. - Quaderni dipartimento economia politica, Università di Siena(2004), pp. 1-24.

Availability:

This version is available at: 2158/676546 since:

Terms of use:

Open Access

La pubblicazione è resa disponibile sotto le norme e i termini della licenza di deposito, secondo quanto stabilito dalla Policy per l'accesso aperto dell'Università degli Studi di Firenze (<https://www.sba.unifi.it/upload/policy-oa-2016-1.pdf>)

Publisher copyright claim:

(Article begins on next page)

QUADERNI



Università degli Studi di Siena
DIPARTIMENTO DI ECONOMIA POLITICA

ALESSANDRO INNOCENTI
MARIA GRAZIA PAZIENZA

n. 438 - Ottobre 2004

ABSTRACT: This paper reports on a laboratory study that examined the influence of experimenter bias in the investment game. Specifically we explored the effect of changing the gender of the experimenter and compared it with the double blind treatment. Our findings show that the presence of a female experimenter influences reciprocity. We also provide further evidence on gender differences in trust and reciprocity.

JEL CODES: C90, C91, J16.

KEYWORDS: experimenter bias, gender differences, trust, reciprocity.

The financial support of the University of Siena and the MIUR is gratefully acknowledged. We thank Francesco Lomagistro for valuable research assistance. Thanks also to the participants in the 2004 International ESA Meeting, Amsterdam, 11–14 June 2004, where a previous version of this paper was presented.

Alessandro Innocenti, Dipartimento di Economia Politica e LabSi, Laboratorio di Economia Sperimentale, Università degli Studi di Siena

Maria Grazia Paziienza, Dipartimento di Studi sullo Stato, Università di Firenze

1. Introduction

An important issue in the methodology of laboratory techniques is experimenter bias.¹ This definition applies to laboratory environments in which experimenters' acts can unconsciously convey to subjects how they should behave in relation to some characteristics of the design and consequently produce biased results. A different case is participant bias, also known as demand characteristics, which applies to experiments in which participants act in ways they believe correspond to what the experimenter is looking for. Thus, if participants modify their spontaneous behaviour to match the real or the presumed aims of the experimenter, results are also biased.

These sources of bias are differentiated by the fact that the former is explicitly related to some specific act or characteristic of the experimenter, while the latter refers generically to the experimental design, but it is not always easy to discriminate between them. What matters more for experimental methodology is that in both cases some features of the laboratory environment may induce subjects to change their choices in order to comply with the experiment's purpose. Once the experimenter becomes aware of this effect, data interpretation has to be revised in order to assess the experimental findings correctly.

The very robust finding that subjects prefer fair to maximized payoffs² has been attributed to the effect of experimenter's observation. Hoffman, McCabe, Shachat and Smith (1994) and Bolton and Zwick (1995) name this effect the anonymity hypothesis and give two reasons for it. The first is related to subjects' participation in future experiments. If the experimenter's presumed aim is to provide evidence against self-seeking behaviour, a subject's preference for fair payoffs could increase his or her probability of being recruited again by the same experimenter. The second reason has a more psychological flavour. Subjects could be concerned with the experimenter's judgement and believe that he or she disapproves of maximized choices.

¹ It is only recently that experimental economists have become aware of this problem. In particular, Hoffman, McCabe and Smith (1996) argue that the subjects' degree of social distance from the experimenter may affect subjects' behaviour, especially determining expectations of reciprocity. Bolton and Zwick (1995) try to find evidence for the hypothesis that experimenter's observation distorts subjects' objectives by testing an ultimatum game.

² Ledyard (1995) and Schram (2000) present extensive surveys documenting that, in a variety of experimental situations, subjects deviate from own-payoff-maximizing behaviour.

Both arguments can be criticized. Most experimenters prefer inexperienced subjects to experienced subjects, in order to control better for learning processes. The beliefs of experimenters about economic or moral principles are not easily predictable by the participants in experiments. However, the importance of the experimental result of subjects deviating from maximizing their own payoffs requires further analysis of this issue.

In this light, Orne (1962) offers another plausible explanation: “The subject's performance in an experiment might almost be conceptualized as problem-solving behavior; that is, at some level he sees it as his task to ascertain the true purpose of the experiment and respond in a manner which will support the hypotheses being tested. Viewed in this light, the totality of cues which convey an experimental hypothesis to the subject become significant determinants of subjects' behavior.” (Orne 1962, p. 778).

Experimenter bias can become relevant because subjects have a real stake in viewing their performance as meaningful. During the experiment, they are constantly wondering about what the experimenter is trying to test and any cue they find useful to answer this question can influence their behaviour. If the design incorporates explicit suggestions as treatment variables, data interpretation is unbiased. Otherwise, expectations about the experimental purpose can affect subjects' behaviour and consequently distort results.

Observable characteristics of the experimenter may be among these cues. Indeed the object of the experimenter can also be inferred from his or her gender. For example, the presence of a female experimenter in tests of gender differences may induce participants to guess that the experiment's purpose is related to discrimination against women. This source of bias can also be reinforced if subjects are informed about the gender of other participants.

Although several recent studies focus on gender effect,³ we are not aware of laboratory experiments that try to analyse the effect of experimenter bias across gender differences.⁴ Our study intends to provide evidence on this issue by testing the

³ See, among the others, the experimental studies on dictator games by Andreoni and Vesterlund (2001) and Eckel and Grossman (1998).

⁴ Andreas Ortmann and Lisa Tichy (1999) tried to deal with the experimenter bias problem by jointly conducting a test on gender differences in prisoner's dilemma. However, the physical presence of two experimenters, one male and one female, does not remove the possibility that one of the two experimenters is perceived by the participants as the determinant one. For

differences between men’s and women’s individual choices in three different treatments. The first and second treatments were conducted by a female and a male experimenter respectively, while the third treatment adopted a double blind procedure. We consider our experiment useful also because it offered additional evidence of gender differences on trust and reciprocity, which is still a controversial issue in economics.

The object of our test was the investment game, also known as the trust game, which was first tested by Berg, Dickhaut and McCabe (1995). Two players are paired off anonymously and respectively named as the sender and the receiver. The sender is given a certain amount of money and told that he or she can keep the entire amount or send some or all of it to the receiver. Any money passed from the sender to the receiver is tripled by the experimenter and then given to the receiver. The receiver can keep the entire amount or give back some or all of it to the sender. When the sender receives the amount sent back by the receiver the game ends. The backward induction solution of the game predicts that the receiver will not send any money back. Anticipating the receiver’s decision, the sender will not send any money to the receiver.

Results from earlier experiments are inconsistent with the conventional game theory prediction. Table 1 offers a summary of the previous work on the investment game.

Table 1. Experimental results on the investment game

	Berg, Dickhaut and McCabe (1995)		Buchan, Croson and Johnson (2000)		Schwieren and Sutter (2003)	Burks, Carpenter and Verhoogen (2003)		Chaudhuri and Gangadharan (2003)	
	No history	Social history	Only U.S.	All		Single role	Both roles, no prior	Both roles, prior	
Trust	51.6	53.6	49	67	65.7	65.0	65.2	47.3	43.3
Reciprocity	30.1	40.2	22	37	37.6	43.6	25.9	17.1	17.5

Trust = Average fraction sent (Amount sent / Initial endowment); Reciprocity = Average fraction returned (Amount sent back/ Amount received)

Notes: In the social history treatment by Berg, Dickhaut and McCabe (1995) subjects were given a summary of the no history treatment results as part of the instructions. Buchan, Croson and Johnson’s (2000) international experiment tested the investment game in China, Japan, Korea and United States. Burks, Carpenter and Verhoogen (2003) compared three treatments: the first (single role) in which subjects played either sender or receiver role, the second (both roles, no prior) in which subjects played both roles but they did not know it before sending, and the third (both roles, prior) where subjects knew that they played both roles before any decisions were made.

example, the person who reads the instructions aloud is usually considered to lead the experiment.

Even if there are significant variations across tests, the relevance of altruistic motivations and intentions to reciprocate was widely confirmed in the laboratory.⁵

Other studies examined gender differences in the investment game. Table 2 presents their results.

Table 2. Experimental results on gender differences in the investment game

	Croson and Buchan (1999)		Chaudhuri and Gangadharan (2001)	
	Men	Women	Men	Women
Trust	69.0	63.0	53.0	34.7
Reciprocity	28.6	37.4	14.7	19.8

Trust = Average fraction sent (Amount sent/Initial endowment); Reciprocity = Average fraction returned (Amount sent back/Amount received)

In both tests men exhibited greater levels of trust and lower levels of reciprocity than women did, even if the difference between men and women was not statistically significant for trust in Croson and Buchan (1999) and reciprocity in Chaudhuri and Gangadharan (2001).

Our experiment adopted nearly the same design as those tests, including that each subject played both roles of sender and of receiver, but we departed from them by imposing the following variant: when participants played as senders, they were informed about the gender of the receiver with whom they were paired off.

The relevance of gender pairing in bilateral relationships was proved experimentally by Sutter, Bosman, Kocher and van Winden (2003). They found that cooperation between players was lower when bargaining partners have the same gender than when they have the opposite gender. In our design, the information about their receiver's gender aimed at focusing senders' attention on the gender variable. In this way, a clear signal about the experimental purpose was conveyed to all subjects. If the experimenter's presumed aim was to study the effect of gender differences on trust and reciprocity, the experimenter's gender might affect subjects' behaviour. For example, the effect of the presence of a female experimenter might be an increase in the degree of trust in women or a change in women's degree of trust in receivers of both sexes. More

⁵ Abbink, Irlenbusch and Renner (2000), Cox (2001) and Cox, Sadiraj and Sadiraj (2002) obtained similar results by testing the moonlighting game. In the moonlighting game, the sender can choose if she wants to give to the receiver part of her endowment or take up to half the endowment given to the receiver. The receiver can decide whether she wants to give or take money from the sender.

generally, we expected subjects in the double blind treatment to exhibit behaviour significantly different from that in the other two treatments, because the signals conveyed by the experimenter's gender were absent. In addition, we expected subjects to change their behaviour significantly in relation to the experimenter's gender. We might conjecture that the presence of the female experimenter could be perceived as relatively unusual and consequently conveys a stronger signal than the presence of a male experimenter about the experimental purpose.

To examine further the difference across treatments we chose not to inform receivers about their sender's gender. This feature of the design was intended to differentiate factors influencing reciprocity and trust. If subjects, when they played as receivers, did not know the gender of their paired sender, their decision to reciprocate could be considered dependent only on the amount of money received and on the experiment's perceived purpose.

To summarize, our experiment intended to verify the effect of experimenter bias by testing two hypotheses. The first hypothesis was that subjects' behaviour in the double blind treatment is significantly different from that in the other two treatments. The second hypothesis was that subjects change their behaviour significantly in relation to experimenter's gender, presumably by increasing the degree of trust and reciprocity in the female experimenter treatment.

The rest of the paper is organized as follows. The experimental design is introduced in Section 2. In Section 3, we report the experimental results, focusing on the effect of experimenter's gender, sender's gender and receiver's gender. Section 4 offers some concluding remarks.

2. Experimental design

The experiments were carried out in the spring of 2004. We submitted the investment game to 94 subjects: 46 women and 48 men. They were undergraduate students in economics from the University of Siena and students in political sciences from the University of Florence, recruited from economics courses through billboards posted around the campuses of the two universities.

The experiments were run manually. The participants were paid according to the euros earned. There was no participation fee.

We ran three treatments. The only difference between the first and the second treatment was that a female and a male respectively played the role of experimenter.⁶ It was made clear to subjects that in these treatments only the experimenter was able to attribute individual choices to individual people. However, anonymity between subjects was guaranteed. The third control treatment adopted a double blind procedure.

Table 3 presents the number of participants for each session and treatment.

Table 3. Number of participants per treatment by gender

Session	Treatment	Participants (female + male)
1	Female experimenter	8 + 6
2	Female experimenter	7 + 9
3	Male experimenter	7 + 9
4	Male experimenter	8 + 8
5	Double blind	8 + 8
6	Double blind	8 + 8
		46 + 48

In the female and male treatments, subjects were first identified by numbers. These numbers were randomly assigned and determined the pairings of senders and receivers. Then each subject was directed to an isolated desk to make his or her decision privately. There they received written instructions. The first part of the instructions was read aloud by the experimenter of the appropriate gender. The second part contained a short questionnaire that was answered at the end of the experiment.

When the experiment began, subjects were given a large unmarked envelope which contained the money to be invested (5 euros which could be transferred in steps of half units), a card marked with the identification number and a small envelope marked with a circle that was either pink or blue. Subjects were asked to remember their numbers. The correspondence between each number and each participant remained unknown to the other participants but not to the experimenter and this was made clear to the participants. Subjects were also informed that if the circle was pink (blue) the person to whom they were to send money was a female (male). In this way, the sender knew the receiver's gender but the sender's gender remained unknown to the receiver.

Senders decided how many euros to keep and how many euros they wish to send to their partner by inserting them in the small envelope. The experimenter collected the small envelopes, recorded the amount sent, tripled it and placed the tripled money into the same envelope for delivery to the appropriate receiver. Receivers then opened their envelopes and decided how much of the money received to return to the sender. The experimenter again collected the envelopes, recorded the amounts returned and gave the envelopes back to the senders. Subjects were informed in the written instructions that they would be playing both roles of sender and receiver but also that the receiver to whom they were paired as sender, would not be their sender when they played the role of receiver.

In the double blind treatment, the design had to guarantee to the participants that the experimenter was unable to attribute individual choices to individual subjects and to avoid any hint about the experimenters' gender. Subjects were gathered in a room where two undergraduate students, a male and a female previously instructed to play the role of monitors, gave them instructions to read privately. When the experiment began, subjects were given a large unmarked envelope, which contained 5 euros, a smaller envelope and a numbered identification card. The small envelope was marked with a pink or a blue circle. As in the previous treatments, subjects were informed that the coloured circle on the small envelope identified the gender of their paired receiver. Moreover, they were asked to remember their numbers. The correspondence between these numbers and identities of the subjects remained unknown to the experimenters, to the monitors and to the other participants at all times and this was made clear to the participants.

Once senders had decided how much money to send to their partners in the small envelopes, they had to insert the identification cards in the smaller envelopes. Moreover, each subject had to write on the identification card the letter F, if female, or the letter M, if male. The sealed envelopes were collected in a closed urn and were taken by the monitor to the experimenters in another room. After recording the amount sent and tripling it, the experimenters marked each larger envelope with the number identifying a receiver of the appropriate gender. The envelopes in the closed urn were

⁶ The authors served as experimenters for the two treatments. In each session, there was also an assistant of the same gender as the experimenter.

delivered by the monitor to the subjects' room. At this time, subjects were called one at a time by the monitor. Once called, a subject had to privately choose the envelope with her or his identification number from the urn placed on an isolated desk. Having decided how much of the money received to return to the sender, subjects sealed their envelopes. The monitors again collected the envelopes, and took them to the experimenters' room, where the experimenters recorded the amounts returned and gave the envelopes back for distribution to senders by the same procedure used before. When the experiment was over, all subjects left the room without revealing their identities.

To summarize, our main variations with respect to the reference design by Berg, Dickhaut and McCabe (1995) were the following ones:

1. All subjects played both roles (sender and receiver) and they knew it before any decisions were made.⁷

2. Senders were informed of their receiver's gender, but receivers did not know their sender's gender.

3. Only the third treatment was double blind, while in the first and in the second treatment, the experimenter was able to attribute individual choices to individual subjects and this was made clear to the participants.

3. Results

Our discussion of the experimental findings addresses the question of whether men and women made different choices across the three treatments, first for trust, and then for reciprocity.

Table 4 presents senders' behaviour in the three treatments across gender.

⁷ Burks, Carpenter and Verhoogen (2003) made the same assumption in their third treatment, which was named as 'both roles, prior' (see Table 1).

Table 4. Trust per treatment by sender's gender

Treatment	Female experimenter		Male experimenter		Double blind		Total	
	Mean	<i>StdD</i>	Mean	<i>StdD</i>	Mean	<i>StdD</i>	Mean	<i>StdD</i>
Sender's gender								
Women	43.3	35.4	39.3	22.2	32.5	23.8	38.3	27.4
Men	28.0	27.0	24.1	18.4	39.4	27.2	30.4	24.8
Total	35.7	31.9	31.3	21.4	35.9	25.4	34.3	26.3

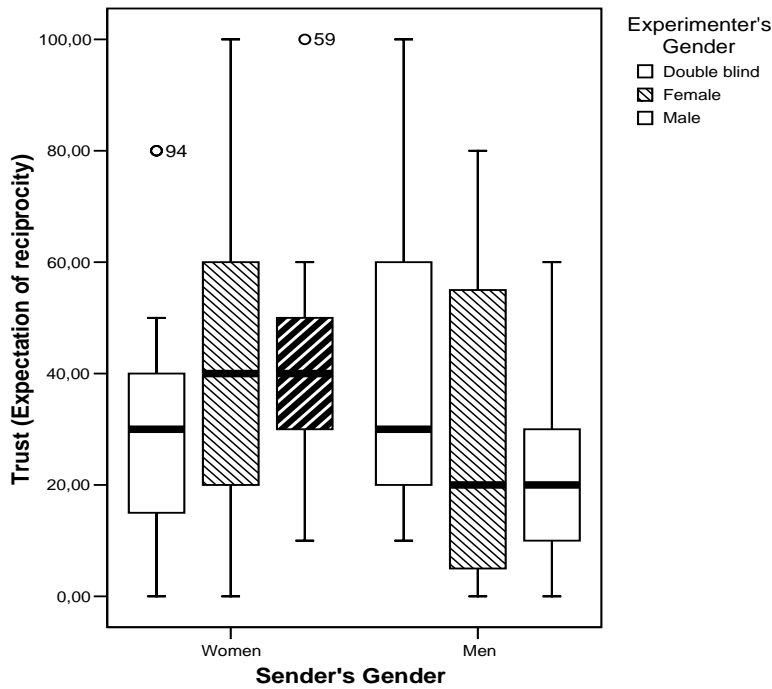
Trust = Average fraction sent (Amount sent/Initial endowment)

The trust indicator is computed as the average ratio between the amount sent and the initial endowment. Contrary to the previous experimental evidence (see Table 2), women showed on average a higher degree of trust than men, but the difference between the two means (38.3 and 30.4) was not significantly different from zero. In the aggregate, trust (34.3) was lower than in the previous tests presented in Table 1. However, the results of the double blind treatment were very similar to that of Chaudhuri and Gangadharan (2001), in particular in the average fraction sent by women (32.5 versus 34.7).

We found similar values of trust across the three treatments (35.7, 31.3, and 35.9). The differences between women and men were not significant as well. Although women trusted less in the double blind treatment (32.5), and men trusted less in the male experimenter treatment (24.1), the difference in the fraction of the amount sent between women and men was not significant using either a t-test (with a t-value of 0.04 and a p-value of 0.97) or a Wilcoxon rank sum test (with a z-value of 911 and a p-value of 0.63).

In Figure 1, data on trust are summarized with box plots. The plots report the median of the data distribution, and the interquartile range to measure the data dispersion. The dispersion of data, which is higher in the female experimenter treatment, explains why the mean differences are not statistically significant. Nevertheless, the figure shows that the difference between men's and women's median values was greater than the difference between the corresponding average values.

Figure 1. Distribution of trust per treatment and sender's gender



In our design, senders were informed of their paired receiver's gender. This information was specifically given to emphasize the effect of experimenter bias. Table 5 presents the average values of trust by receiver's gender. In the aggregate, senders trusted more men (35 %) than women (33%), but the difference between the two average values was not significant and the medians were nearly identical (Figure 2). In the male and in the female experimenter treatments, senders exhibited a higher degree of trust in men than in women, while in the double blind treatment the reverse was true. There was no evident gender pairing effect because the degree of trust was quite similar across pairs of the same gender and of opposite gender.

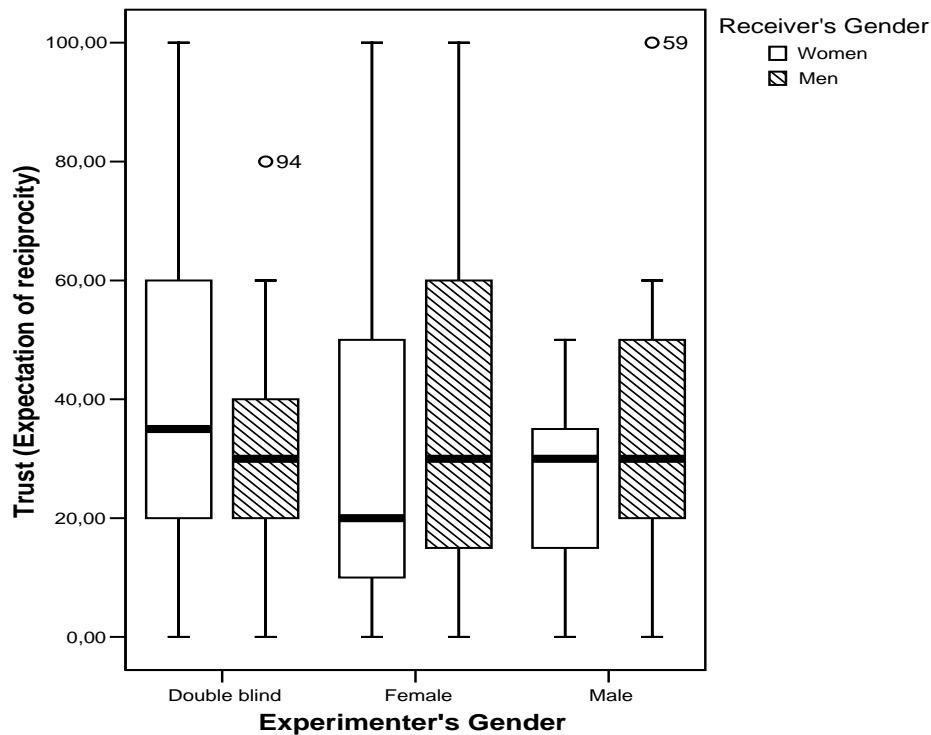
Table 5. Trust per treatment by receiver's gender

		Treatment			
		Female experimenter	Male experimenter	Double blind	Table total
Total (N=94)					
Receiver's gender	Women	30.67	26.67	41.88	33.26
	Men	40.67	35.29	30.00	35.21
	Total	35.67	31.25	35.94	34.26
Male senders (N=48)					
Receiver's gender	Women	21.25	21.00	50.00	30.00
	Men	35.71	28.57	28.75	30.91
	Total	28.00	24.12	39.38	30.42
Female senders (N=46)					
Receiver's gender	Women	41.43	38.00	33.75	37.50
	Men	45.00	40.00	31.25	38.85
	Total	43.33	39.33	32.50	38.26

Trust = Average fraction sent (Amount sent/Initial endowment)

The box plots depicted in Figure 2 show that the amount of money sent to both men and women was more dispersed in the female experimenter treatment than in the other two treatments.

Figure 2. Distribution of trust per treatment and receiver's gender



Apart from this distinction, our findings show that there was no systematic difference in trust indicator among treatments. Experimenter gender did not seem to significantly influence senders' choices.

Next, we turn to the analysis of receivers' choices. Table 6 presents an indicator of reciprocity, i.e. the average fractions returned by receivers.

Table 6. Reciprocity per treatment by receiver's gender

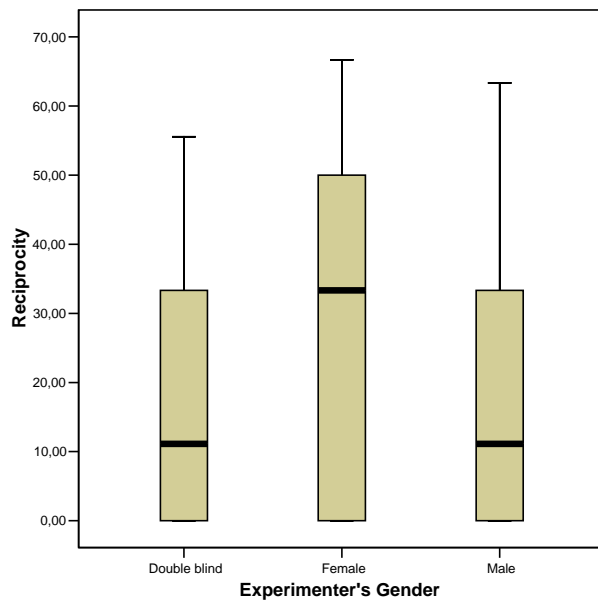
Treatment	Female experimenter		Male experimenter		Double blind		Total	
Receiver's gender	Mean	<i>StdD</i>	Mean	<i>StdD</i>	Mean	<i>StdD</i>	Mean	<i>StdD</i>
Women	27.3	27.0	21.0	17.9	20.4	20.8	22.6	21.5
Men	30.9	22.6	9.7	17.7	18.9	19.4	19.0	21.1
Total	29.1	24.4	15.2	18.4	19.7	19.8	20.8	21.3

Reciprocity = Average fraction returned (Amount sent back/Amount received)

The stronger pattern of behaviour was that the presence of a female experimenter induced subjects of both genders to reciprocate more than in the other treatments. This result is also shown in Figure 3,⁸ which presents the median values and the interquartile range of the distribution in the three treatments.

⁸ The choice of the proportion of the amount sent back by each receiver as the indicator of reciprocity leads to some missing cases. If the receiver was sent nothing by the paired sender, he or she is excluded from computation as a missing variable. Otherwise, if the computed value is zero this means that the receiver gives back nothing of what he receives.

Figure 3. Distribution of the average values of reciprocity per treatment



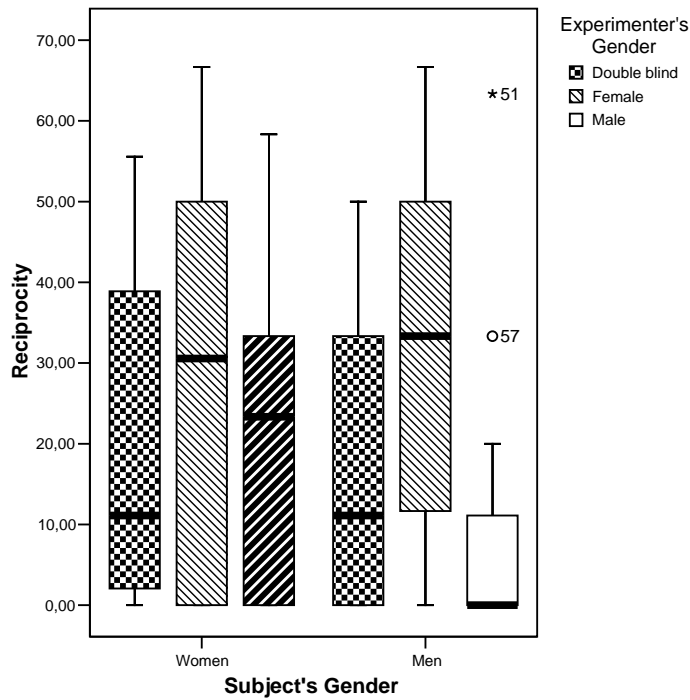
The median values in the male experimenter treatment and in the double blind treatment were quite similar and significantly lower than in the female experimenter treatment. Statistical tests also corroborate this difference for the average values. Table 7 shows that both the t-test and the Wilcoxon test were significant at 95%.

Table 7. Statistical tests for reciprocity among treatments

t test for equality of means	N	Mean	Mean diff.	T	Sig. (2-tailed)
Female experimenter treatment	24	29.12			
Male experimenter and double blind treatments	59	17.44	-11.67	-2.32	0.02
Wilcoxon test					
Wilcoxon W	2289.5				
Z	-1.95				
Sig. (2-tailed)	0.052				

The box plots in Figure 4 show how male receivers particularly showed a lower degree of reciprocity in the male experimenter treatment. This effect is reinforced by the fact that the median value is zero.

Figure 4. Distribution of reciprocity per treatment by subject's gender



The propensity to be “fairer” in the female experimenter treatment is also confirmed by the statistical analysis of men’s behavior.⁹ Table 8 shows that the presence of a female experimenter induced male subjects to reciprocate significantly more than in the other two treatments, using either a t-test or a Wilcoxon rank sum test.

Table 8. Statistical tests for reciprocity among treatments (men only)

t test for equality of means	N	Mean	Mean diff.	T	Sig. (2-tailed)
Female experimenter Treatment	12	30.92			
Male experimenter and double blind treatments	30	14.30	-16.62	-2.44	0.02
Wilcoxon test					
Wilcoxon W	563.5				
Z	-2.35				
Sig. (2-tailed)	0.02				

⁹ The amount of money received from the paired sender and the percentage of money sent back to the paired sender was positively correlated with the Spearman’s correlation coefficient of 0.46 (0.576 for women and 0.372 for male), which was significant at the 0.01 level in all cases.

These outcomes may be interpreted as evidence of experimental bias for reciprocity. The presence of the female experimenter could have influenced subjects' expectations and changed their spontaneous behaviour. According to our proposed justification for experimenter bias, receivers increased the degree of reciprocity to fulfil the perceived experimental purpose that was presumably the testing of trustworthiness. But experimenter bias would have been effective only for the treatment in which the signal used to infer this purpose was perceived by subjects as more evident, specifically the presence of a female experimenter. A rationale for this outcome is that the subjects perceived the female experimenter as unusual. The effect was to increase the degree of reciprocity behaviour in all subjects, male and female. The same effect did not emerge for trust. An explanation may be found in the difference between the factors influencing trust and reciprocity, as we discuss in the next section, which draws some conclusions.

4. Conclusions

Our experiment intended to test the effect of experimenter bias in the laboratory. Among the possible features of an experimental environment eligible for this purpose, experimenter's gender was a straightforward option, especially if the design adopts gender differences as a treatment variable. Moreover, we chose to test the investment game that has been widely investigated in the last decade. Another reason to opt for the investment game was that two different kinds of behaviour could be studied in the same experiment, i.e. trust and reciprocity. Even in the one-shot version of the game, these two tendencies can be the product of different motivational traits. Trust can be perceived as an investment and therefore be dependent mainly on risk aversion; reciprocity or trustworthiness is usually more affected by psychological or ethical factors, which range from the perceived kindness of trusting to inequality aversion. This distinction motivated us in defining both theoretical premises and experimental design. The theoretical starting point was the conjecture that experimental subjects try to discover the true purpose of the experiment and may change their spontaneous behaviour to confirm the experimenter's expectations. Any signal useful to perform this task can influence their behaviour. The gender of the experimenter may be the source of

a bias if gender differences are explicitly considered as variables determining choices. The experimental design took into account the distinction between trust and reciprocity by differentiating subjects' information: trusters were informed of their paired trustee's gender but trustees did not know their paired truster's gender. We observed behaviour in three different treatments: the female experimenter treatment and the male experimenter treatment, conducted respectively by the female and male authors of this paper, and the double blind treatment, where complete anonymity between subjects and between subjects and experimenter was guaranteed. The proposed interpretation of experimenter bias led to the prediction that subjects would modify their behaviour in relation to the experimenter's gender and to the way it is related to the experimental purpose. The analysis of the data indicates that (i) there was no significant evidence that experimenter's gender influenced the degree of trust; (ii) there was statistically relevant evidence that the presence of a female experimenter induced a higher degree of reciprocity than in other treatments; (iii) this effect concerned all subjects, regardless of their gender.

We interpret these findings as generally supportive of the relevance of experimenter bias. If the cues provided by laboratory environment are clear enough to convey a specific experimental purpose, as the presence of a female experimenter in a test of gender effect did, subjects' behaviour is significantly affected. In our experiment, trust was unbiased presumably because it was perceived as a behavioural pattern dependent on risk attitude and self-seeking behaviour. On the contrary, the tendency to reciprocate was perceived as dependent on considerations of fairness that were associated by the subjects with the presence of the female experimenter.

Finally, our study provides new evidence on gender differences in trust and reciprocity. The degree of trust and reciprocity was generally lower than in previous experiments. Women trust and reciprocate more than men, but while the result for reciprocity is in keeping with previous studies that for trust is not. We also found that gender differences were higher for trust than for reciprocity, although not statistically significant. Our results corroborate the view that gender effect depends more on the details of the experimental structure than on behavioural differences among men and women.

A number of questions are left unanswered by our study. First, is the bias produced by the presence of the female experimenter generally true or does it vary between countries? Second, how is it possible to manipulate an experimental environment to nullify gender effect in the double blind treatment? Third, are experiments on gender effect necessarily biased when subjects are informed of other subjects' genders? We consider these issues relevant for the study of gender differences in the laboratory and we intend to address them in future experiments.

References

- Abbinck, Klaus, Bernd Irlenbusch and Elke Renner, (2000), "The Moonlighting Game: An Empirical Study on Reciprocity and Retribution." *Journal of Economic Behavior and Organization*, 42, pp. 265–77.
- Andreoni, James, and Lisa Vesterlund, (2001), "Which is the Fairer Sex? Gender Differences in Altruism." *Quarterly Journal of Economics*, 115, pp. 293–311.
- Bacharach, Michael, Gerardo Guerra and Daniel Zizzo, (2001), "Is Trust Self-Fulfilling? An Experimental Study." Discussion paper, Department of Economics, University of Oxford.
- Berg, Joyce, John Dickhaut and Kevin McCabe, (1995), "Trust, Reciprocity, and Social History." *Games and Economic Behavior*, 10, pp. 122–42.
- Bolton, Gary E. and Rami Zwick, (1995), "Anonymity versus Punishment in Ultimatum Bargaining." *Games and Economic Behavior*, 10, pp. 95–121.
- Buchan, Nancy R., Rachel T.A. Croson and Eric J. Johnson, (2000), "Trust and Reciprocity: An International Experiment." Discussion paper, University of Wisconsin-Madison, University of Pennsylvania and Columbia University.
- Burks, Steven, Jeffrey Carpenter and Eric Verhoogen, (2003), "Playing both roles in the trust game." *Journal of Economic Behavior and Organization*, 51, pp. 195–216.
- Chaudhuri, Ananish and Lata Gangadharan, (2001), "Gender Differences in Trust and Reciprocity." Discussion paper, Wellesley College and University of Melbourne.
- Chaudhuri, Ananish and Lata Gangadharan, (2003), "Sending Money in the Trust Game: Trust or Other Regarding Preferences." Discussion paper, Wellesley College and University of Melbourne.

- Cox, James C., (2001), "On the Economics of Reciprocity." Discussion paper, University of Arizona.
- Cox, James C., Klarita Sadiraj and Vjollca Sadiraj, (2002), "Trust, Fear, Reciprocity, and Altruism." Discussion paper, University of Arizona and University of Amsterdam.
- Croson, Rachel, and Nancy Buchan, (1999), "Gender and Culture: International Experimental Evidence from Trust Games." *American Economic Review*, 89, pp. 386–91.
- Eckel, Catherine C., and Philip J. Grossman, (1998), "Are Women Less Selfish than Men: Evidence from Dictator Experiments." *The Economic Journal*, 108, pp. 726–35.
- Hoffman, Elizabeth, Kevin McCabe, Keith Shachat and Vernon L. Smith, (1994), "Preferences, Property Rights, and Anonymity in Bargaining Games." *Games and Economic Behavior*, 7, pp. 346–80.
- Hoffman, Elizabeth, Kevin McCabe and Vernon L. Smith, (1996), "Social Distance and Other-Regarding Behavior in Dictator Games." *American Economic Review*, 86, pp. 653–60.
- Ledyard, John O., (1995), "Public goods: a survey of experimental research", in *The Handbook of Experimental Economics*, edited by J. Kagel and A. Roth, Princeton: Princeton University Press.
- Orne, Martin T., (1962), "On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications." *American Psychologist*, 17, pp. 776–83.
- Ortmann, Andreas, and Lisa K. Tichy, (1999), "Gender differences in the laboratory: evidence from prisoner's dilemma games." *Journal of Economic Behavior and Organization*, 39, pp. 327–39.
- Schram, Arthur, (2000), "Sorting Out the Seeking: Rents and Individual Motivation." *Public Choice*, 103, pp. 231–58.
- Schwieren, Christiane, and Matthias Sutter, (2003), "Trust in cooperation or ability? An experimental study on gender differences." Discussion paper, University of Innsbruck.
- Sutter, Matthias, Ronald Bosman, Martin Kocher and Frans van Winden, (2003), "Experimental evidence of the importance of gender pairing in bargaining." Discussion paper, Max Planck Institute, University of Innsbruck, De Nederlandsche Bank and CREED/ University of Amsterdam.

Appendix

Instructions for all participants

A. FEMALE AND MALE EXPERIMENTER TREATMENTS

This is an experiment in the economics of decision-making. The Ministry of University and the University of Siena have provided funds to conduct this research. The instructions you are about to read are self-explanatory. If you follow them closely and make appropriate decisions, you can earn an amount of money that will be given to you in cash at the end of the experiment. If you have any questions, please raise your hand and the experimenter will come to you and answer your question.

In this experiment, each of you will be paired with a different person. You will not be told who these people are either during or after the experiment nor will they be told who the others are. The only information you will have is the gender of the person to which you are paired. The experimenter will be in charge of the envelopes as explained below. In addition, he will verify that the instructions have been followed as they appear here.

Each person will be given 5 euros as a show-up fee for this experiment. Each person will have the opportunity to send in an envelope, some, all or none of their show-up fee to the person which you are paired to. The amount sent will be tripled. For example, if you send an envelope that contains 2 euros, the envelope will contain 6 euros when it reaches the paired person. If you send an envelope that contains 4 euros, the envelope will contain 12 euros when it reaches the paired person. The paired person will then decide how much money to send back to you and how much money to keep.

Each person will play both roles in the experiment. Each of you will be paired with two people. In one pair, you will be the person who decides how many of the 5-euro show-up fee to send to another person, who receives the amount sent tripled by the experimenter. In the other pair, you will be the person who receives the amount sent by another person and tripled by the experimenter and you will decide to send back some, all, or none of the amount received to the sender. So each of you will take two decisions. However, the important thing to bear in mind is that you are not paired with the same person as sender and receiver. Rather you are paired with two different people.

The remainder of these instructions will explain exactly how the experiment is run. This experiment is structured so that no one, except the experimenter, will know the personal decision of people. Since your decision is private, we ask that you do not tell anyone your decision during, or after, the experiment.

The experiment is conducted as follows: a number of large unmarked envelopes have been placed in a box. Each of these envelopes contains 5 euros as a show-up fee for this experiment, a card marked with an identification number that you are asked to remember, and a smaller envelope marked with a circle, which will be coloured pink or blue. The experimenter will hand one person at a time an unmarked envelope from the box. Once a person has an envelope, he or she will privately open the unmarked envelope and place as many euros in the smaller circled envelope as they want, keeping the rest. Examples: (1) put 2 euros in the smaller envelope and keep 3 euros; (2) put 4 euros in the smaller envelope and keep 1 euro. These are examples only; the actual decision is up to each person.

It is important to keep in mind that the person who received the amount you sent will be a female if the smaller envelope is marked with a pink circle and he will be a male if the smaller envelope is marked with a blue circle. This process will continue until everyone has made his or her decision.

Once a person has made a decision, the experimenter will collect all the larger, unmarked envelopes, and return them to the box. Notice that each returned envelope will look exactly the same.

After all the envelopes have been put in the return box, the experimenter will then privately, one at a time, take the smaller envelopes out of the larger envelopes, record on a sheet of paper the number of the identification card and the amount of money inside the smaller envelope. The experimenter will then triple the amount of money in the smaller envelope and place the smaller envelope back into the larger envelope. At this point, the experimenter will transfer the envelopes in the return box.

The experimenter will then give to each person, one at a time, an unmarked envelope from the box. Each of you will privately open the larger envelope and must decide how many euros to leave in the smaller envelope. The person keeps the remaining euros. The smaller envelope should then be placed in the larger envelope. When everyone has had the opportunity to make his or her decision, the experimenter will collect the larger envelopes and return them to the box. The experimenter will then privately, one at a time, open the larger envelopes and record how much is in the smaller envelope. After recording how much was in the smaller envelope, the experimenter will put the smaller envelope in the larger envelope, and will place them back in the return box.

Then experimenter will choose one person at a time to go to the box marked return envelopes to retrieve the smaller envelope with the appropriate identification number marked on it. Do not open your envelope yet. This process will continue until everyone has retrieved his or her envelope and returned to his or her seat. When everyone is finished, experimenter will ask if everyone has retrieved the correct envelope. If the people have all taken the correct envelope, then the experiment is finished. If, however, an envelope has ended up with the wrong person, then the experimenter will collect all the smaller envelopes again and the process will repeat until everyone has the correct envelope.

Before leaving the room you, everyone will be asked to fill out a short questionnaire. At the top of the questionnaire, you will be asked for the card identification number. Please do not forget to include this information. Once you have finished the questionnaire, you will be asked to put it in the box placed at the back of the room.

Please raise your hand if you have any questions regarding how the experiment will proceed.

B. DOUBLE BLIND TREATMENT

This is an experiment in the economics of decision-making. The Ministry of University and the University of Siena have provided funds to conduct this research. The instructions you are about to read are self-explanatory. Two of us have been chosen as monitors and will check that the instructions have been followed as they appear here. However, they will not answer any questions during this experiment. If you have any doubts, you should read back through these instructions. Now that the experiment has begun, we ask that you do not talk, at all, during this experiment. If you follow these instructions closely and make appropriate decisions, you can earn an amount of money that will be given to you in cash at the end of the experiment.

In this experiment, each of you will be paired with a different person. You will not be told who these people are either during or after the experiment nor will they be told who the others are. The only information you will have is the gender of the person to which you are paired.

Each person will be given 5 euros as a show-up fee for this experiment. Each person will have the opportunity to send in an envelope, some, all or none of their show-up fee to the person which you are paired to. The amount sent will be tripled. For example, if you send an envelope that contains 2 euros, the envelope will contain 6 euros when it reaches the paired person. If you send an envelope that contains 4 euros, the envelope will contain 12 euros when it reaches the paired person. The paired person will then decide how much money to send back to you and how much money to keep.

Each person will play both roles in the experiment. Each of you will be paired with two people. In one pair, you will be the person who decides how much of the 5-euro show-up fee to

send to another person, who receives the amount sent tripled by the monitors. In the other pair, you will be the person who receives the amount sent by another person and tripled by the experimenter and you will decide to send back some, all, or none of the amount received to the sender. So each of you will take two decisions. However, the important thing to bear in mind is that you are not paired with the same person as sender and receiver. Rather you are paired with two different people.

The remainder of these instructions will explain exactly how the experiment is run. This experiment is structured so that no one, including the experimenters and the monitors, will know the personal decision of people. Since your decision is absolutely private, we ask that you do not tell anyone your decision during, or after, the experiment.

The experiment is conducted as follows: a number of large unmarked envelopes have been placed in a box. Each of these envelopes contains 5 euros as a show-up fee for this experiment, a card marked with an identification number that you are asked to remember, and a smaller envelope marked with a circle, which will be coloured pink or blue. Then monitors will call one person at a time to go to the isolated box placed in the front of the room. Each person will take an unmarked envelope from the box and will come back to his or her isolated desk.

Once a person has an envelope, he or she will privately open the unmarked envelope and write on the identification card the letter F if he is a female or the letter M if he is a male. Please do not forget to include this information. Then each person place as many euros in the smaller circled envelope as they want, keeping the rest. Examples: (1) put 2 euros in the smaller envelope and keep 3 euros; (2) put 4 euros in the smaller envelope and keep 1 euro. These are examples only; the actual decision is up to each person. It is important to keep in mind that the person who received the amount you sent will be a female if the smaller envelope is marked with a pink circle and a male if the smaller envelope is marked with a blue circle. This process will continue until everyone has made his or her decision.

Once a person has made a decision, he or she will put the smaller envelope, and the identification card in the larger envelope. Then the monitors will call one person at a time to go to the isolated box. Each person will put the larger envelope into the box. Notice that each returned envelope will look exactly the same and neither monitors nor others will be able to attribute individual choices to individual subjects.

After all the envelopes have been put in the return box, the monitors will then privately, one at a time, take the smaller envelopes out of the larger envelopes, record on a sheet of paper the letter and the number written on the identification card and the amount of money inside the smaller envelope. The monitors will then triple the amount of money in the smaller envelope, place the smaller envelope back into the larger envelope, and write an identification number on the larger envelope. At this point, the monitors will transfer the envelopes in the return box.

The monitors will then call one person at a time to go to the isolated box to retrieve the larger envelope with his or her identification number marked on it. Do not open your envelope yet. This process will continue until everyone has retrieved his or her appropriate envelope and returned to his or her seat. When everyone is finished, monitors will ask if everyone has retrieved the correct envelope. If the people have all taken the correct envelope, then the experiment will continue. If, however, an envelope has ended up with the wrong person, then the monitors will call one person at a time again and the process will repeat until everyone has the correct envelope.

Then each of you will privately open the larger envelope and must decide how many euros to leave in the smaller envelope. The person keeps the remaining euros. The smaller envelope should then be placed again in the larger envelope. When everyone has had the opportunity to make his or her decision, the monitors will call again one person at a time. Each person will return the larger envelopes to the box. The monitors will then privately, one at a time, open the larger envelopes and record how much is in the smaller envelope. After recording how much was in the smaller envelope, the monitors will put the smaller envelope in the larger envelope, and will place them back in the return box.

Then monitors will call one person at a time to go to the box marked return envelopes to retrieve the smaller envelope with the appropriate identification number marked on it. Do not open your envelope yet. This process will continue until everyone has retrieved his or her envelope and returned to his or her seat. When everyone is finished, monitors will ask if everyone has retrieved the correct envelope. If the people have all taken the correct envelope, then the experiment is finished. If, however, an envelope has ended up with the wrong person, then the monitors will collect all the smaller envelopes again and the process will repeat until everyone has the correct envelope.

At this time, you should take all your belongings and leave the building when you are done. When everyone in the room has left, the experiment is over, and the monitors will be paid for their participation.