

# Cardiff Economics Working Papers



Working Paper No. E2021/3

## Lab-Sophistication: Does Repeated Participation in Laboratory Experiments Affect Pro-Social Behaviour?

*Tiziana Medda, Vittorio Pelligra and  
Tommaso Reggiani*

February 2021

ISSN 1749-6010

Cardiff Business School  
Cardiff University  
Colum Drive  
Cardiff CF10 3EU  
United Kingdom  
t: +44 (0)29 2087 4000  
f: +44 (0)29 2087 4419  
[business.cardiff.ac.uk](http://business.cardiff.ac.uk)

This working paper is produced for discussion purpose only. These working papers are expected to be published in due course, in revised form, and should not be quoted or cited without the author's written permission.

Cardiff Economics Working Papers are available online from:

<http://econpapers.repec.org/paper/cdfwpaper/> and

[business.cardiff.ac.uk/research/academic-sections/economics/working-papers](http://business.cardiff.ac.uk/research/academic-sections/economics/working-papers)

Enquiries: [EconWP@cardiff.ac.uk](mailto:EconWP@cardiff.ac.uk)

---

# Lab-Sophistication: Does Repeated Participation in Laboratory Experiments Affect Pro-Social Behaviour?

Tiziana Medda\*  
Vittorio Pelligra†  
Tommaso Reggiani‡

February 23, 2020

## Abstract

Experimental social scientists working at research-intensive institutions deal inevitably with subjects who have most likely participated in previous experiments. It is an important methodological question to know whether participants that have acquired a high level of lab-sophistication show altered pro-social behavioral patterns. In this paper, we focus both on the potential effect of the subjects' lab-sophistication, and on the role of the knowledge about the level of lab-sophistication of the other participants. Our main findings show that while lab-sophistication per se does not significantly affect pro-social behaviour, for sophisticated subjects the knowledge about the counterpart's level of (un)sophistication may systematically alter their choices. This result should induce caution among experimenters about whether, in their settings, information about lab-sophistication can be inferred by the participants, due to the characteristics of the recruitment mechanisms, the management of the experimental sessions or to other contextual clues.

**Keywords:** Lab-sophistication; Experimental Methodology; External Validity; Pro-social behaviour; Cooperation

**JEL Codes:** D03, D83, C91, C92.

Final version published in *GAMES – special issue on Pro-sociality and Cooperation*.

Open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Medda T, Pelligra V, Reggiani T. (2021). Lab-Sophistication: Does Repeated Participation in Laboratory Experiments Affect Pro-Social Behaviour? *Games*.12(1):18. <https://doi.org/10.3390/g12010018>

---

\* T. M.: University of Turin; [tiziana.medda@unito.it](mailto:tiziana.medda@unito.it)

† V. P.: University of Cagliari, CRENoS; [pelligra@unica.it](mailto:pelligra@unica.it)

‡ T. R.: Cardiff University, Masaryk University, IZA; [reggianit@cardiff.ac.uk](mailto:reggianit@cardiff.ac.uk)  
(Masaryk University - grant number MUNI/G/0985/2017)

## 1. Introduction

Since its first appearance, experimental analysis of economic behaviour has provoked sceptical reactions and criticism [1]. The methodological rigour has steadily increased among the community of experimentalists as an antidote to this criticism. In fact, “Just as we need to use clean test tubes in chemistry experiments, so we need to get the laboratory conditions right when testing economic theory” ([2], F17). One of the most critical conditions refers to the characteristics of the subjects with whom the experiments are run. They can, in fact, affect the generalizability of the experimental results in various ways [3, 4]. In particular, a potential weakness derives from the predominant use of students as experimental subjects. The extensive use of pools of students, in fact, may generate problems related to their intrinsic characteristics (the majority of them are “WEIRD” college students,<sup>1</sup> coming from western industrialized, rich and democratic countries<sup>2</sup>), but the voluntary basis of the enrolment process may also produce self-selection that in turn may lead to the formation of experimental pools with peculiar characteristics.<sup>3, 4</sup>

A further potential source of modification arises from the combined effect of “location” and “number”: given the prevalent location of labs in university campuses<sup>5</sup> and the ever-increasing number of experiments run in each of these labs, students tend to accumulate game-specific and laboratory sophistication through repeated participation in experimental sessions [22].

We investigate whether familiarity with the laboratory context, “lab-sophistication”, given by a long record of participations in laboratory-based studies, alters the subjects’ behaviour in a set of experimental games used to capture pro-social behaviour.

More precisely, by lab-sophistication, we do not mean the experience accumulated in a specific game, but the general familiarity with the standard procedures that characterise laboratory sessions and the framing of the decision situation usually implemented in behavioural experiments [23, 24]. We examine whether having taken part in several experimental sessions - having gained familiarity with the lab context and certain knowledge about the working of an experimental session - generates systematic variations in the pro-social choices made by individuals.

To distinguish subjects in term of their lab-sophistication, we adopt the following criteria: The highly sophisticated subjects (H-types) are subjects with 15 or more participations in lab experiments; the unsophisticated subjects with low laboratory context familiarity (L-types) are those whose number of participations in experiments is between 1 and 5.

Moreover, we explore a closely related issue: whether knowledge about the counterpart’s lab-sophistication plays a systematic role in shaping the behaviour of experimental subjects in simple games focusing on cooperation and pro-sociality. Even though experimenters do not provide explicitly information about the lab-sophistication of their participants neither before nor during the experiment, due to peculiarities of the recruiting system (e.g., small pool) or the management of the experimental sessions (e.g., waiting times outside the lab and socialization of the participants), it cannot be excluded that subjects infer lab-sophistication level of other subjects.

---

<sup>1</sup> See [5-15].

<sup>2</sup>[16].

<sup>3</sup> See [17-20], among others.

<sup>4</sup> Guillen et al. [21] show how males and well performing subjects (in monetary terms) are more likely to return in lab and this may introduce a bias to the conclusions derived from observing their behavior.

<sup>5</sup> According to a list drafted by the Laboratoire d’Économie Expérimentale de Nice, only two out of 166 experimental-economics labs in the world are not located on a university campus, and only one is independent and not related to academic activities (<https://orsee.unice.fr/public/labs.php/>, last accessed on 5 March 2020).

In this context becomes crucial to ask whether the knowledge about the other players' level of lab-sophistication potentially plays a role in modifying subjects' choices. Does knowing that the player with whom you are interacting is a lab-sophisticated (lab-unsophisticated) player affect your choices? And if you are a lab-sophisticated (lab-unsophisticated) player, does the knowledge that the player with whom you are interacting is a lab-sophisticated (lab-unsophisticated) player affect your choices?

Thus, in this paper we focus on the role of subjects' lab-sophistication and on its consequences on pro-social behaviour.

Our results show that, while lab-sophistication *per se* does not significantly affect the behaviour of participants, information about the counterpart's level of lab-sophistication may systematically alter the choices of those, among the participants, that are already lab-sophisticated, especially in the situations like the ultimatum game and the trust game. We do not have a theoretical explanation for this phenomenon which, however, we think represents an important practical and methodological element that needs to be considered when i) recruiting subjects [25], ii) designing, iii) running a lab experiment, iv) and when interpreting or assessing its conclusions.

The remainder of the paper is organised as follows. In Section 2 we provide a brief overview of the literature and outline our contribution with respect to the studies that most closely relates to the present paper. In Section 3, the experimental design and procedures are described. In Section 4, the testable hypotheses are derived. Section 5 describes the statistical analyses and the main results. Section 6 concludes.

## 2. Related literature

The need to clarify the role of laboratory experience in economic experiments has given rise, in the last few decades, to a small, but growing, stream of studies: Harrison et al. [26] and Benson and Faminov [27] discuss IO experiments; Marwell and Ames [28], Isaac et al. [29] and Bolton [30] focus on bargaining-game experiments; and, more recently, Matthey and Regner [31], Capraro and Cococcioni [32], Xue et al. [33], Benndorf et al. [34] and Conte et al. [35] have published on this subject.

Among the more recent contributions, Conte et al. [35] focus on a public good game and show that subjects' contribution behaviour is affected by the number of previous participations. However, the effect is stronger if subjects have gained such experience taking part in experiments involving the same typology of game compared to the experience gained in experiments implementing different games. Focusing on cooperation in one-shot interaction, Capraro and Cococcioni [32] analyse the history-dependent dynamic process. Authors run a standard two-person Prisoner's Dilemma in which participants are randomly assigned to either of two conditions: (i) Time pressure condition, which measures intuitive cooperation, and (ii) time delay condition, which measures deliberate cooperation. They report that promoting intuition versus deliberation, has no effect on cooperative behaviour among inexperienced subjects playing in a non-cooperative setting, and that experienced subjects cooperate more than inexperienced subjects, but only under time pressure.

Xue et al. [33] test whether participants with greater experience in experiments will have higher maximization rates in prepayment treatments. Their results show that, individuals who participated in many economics experiments, do not choose differently to those who are novices.

Matthey and Regner [31] and Benndorf et al. [34] provide the two contributions that are closely related to ours. In the former, the authors consider data from three different studies involving ‘allocation games’; their results show that subjects with a higher number of participations tend to be less generous, but this holds true only for participation in experiments with similar tasks.

Benndorf et al. [34] main result is a null result, as they find no ‘experience effect’ for all the games they consider, with the exception of the trust game, where they observe less trust and trustworthiness from the experienced participants. Moreover, they find a moderate, but significant, ‘recruitment bias’, in that experienced subjects tend to sign up for sessions more often than inexperienced ones.

Our study contributes to the existing literature in several ways: first, we do not restrict our scope to the effect of ‘game-specific experience’, as in [31, 35]; we, in the contrary, consider a multidimensional set of games to explore the broader effect generated by ‘lab-sophistication’ intended as a certain degree of familiarity with laboratory’s procedures and settings. Second, our contribution differs from the previous ones, as far as the quantitative definition of ‘experience’ is concerned. We denote, in fact, the subjects in our pool as lab-sophisticated only if they have a considerably large number of lab participations (at least fifteen). Matthey and Regner [31] consider subjects with at most thirteen participations; Conte et al. [35] define as ‘experienced’, subjects with, on average, 2.2 participations in public-goods game and, 5.2 participations in experiments involving different games. The paper by Benndorf et al. [34] is, in this respect, the closest to ours, as they consider ‘experienced’ subjects with at least ten participations.

However, the most important difference between our study and the previous ones concerns the role of information. From a methodological point of view this is a relevant issue every time we study strategic interactions.

### 3. Experimental Design and Procedures

Our main goal is twofold. First, we investigate the effect of high versus low level of lab-sophistication in decision-making experiments: we do this considering a set of simple experimental games where fairness and reciprocity principles are concerned. Secondly, we study the effect of the information about the counterpart’s lab-sophistication level on players’ choices.

Exploiting data stored in the ORSEE recruitment system [36] of the University of Cologne, we recruited, by design, participants with a very different number of participations in previous experiments. We consider as lab-sophisticated subjects (**H-types**) individuals with at least 15 previous participations in experiments, and as lab-unsophisticated subjects (**L-types**) individuals with 1 to 5 participations.<sup>6</sup> These two pools have been chosen in order to assure (i) an adequate number of subjects in each group and (ii) a sharp difference in the level of individual laboratory sophistication between the two groups.

In our first experimental condition (**C1**), each participant is asked to make her decisions in four standard experimental games, without receiving any information about the level of lab-

---

<sup>6</sup> We excluded from the recruiting phase the cluster of registered subjects having 0 participations simply because in relative terms this group represents the less populated one (with the seasonal exception represented by the period in which freshers start their courses).

sophistication of the counterpart<sup>7</sup> with whom she is randomly matched. Here we investigate the effect of having acquired high lab-sophistication (H-types) vs. low lab-sophistication (L-types), merely through repeated participations.

Moreover, we consider a within-subject design<sup>8</sup> in which, in each session, each subject plays the same four standard games faced in C1, under the following conditions:

*Condition 2 (C2)*: each participant is asked to make decisions in the four games knowing that she is now randomly re-matched with a different subject having the same level of lab-sophistication: H-type vs. H-type / L-type vs. L-type;

*Condition 3 (C3)*: each participant is asked to make decisions in the four games knowing that she is now randomly re-matched with a different subject (with respect to C1 and C2) having a different level of lab-sophistication: H-type vs. L-type / L-type vs. H-type (see Table 1).

The baseline condition **C1** was always played as the first in all six sessions, while the ordering of the other two experimental conditions was randomized.<sup>9</sup> In **C1**, subjects were randomly paired and no further information was given to them. In **C2** and **C3**, we revealed the following information in the instructions:

**C2**: for H-type [L-type]: "...in this situation, you will face a different counterpart who has a HIGH [LOW] level of experience. That is, a subject who has participated in many [in few] experiments";

**C3**: for H-type [L-type]: "...in this situation, you will face a different counterpart who has a LOW [HIGH] level of experience. That is, a subject who has participated in few [in many] experiments".

Each session, by design, was balanced in terms of the number of H and L types. While in **C1**, types were paired randomly, in **C2** and **C3** we used the ad-hoc assortative matching rule, randomly pairing, within the same session, **H** (L) with **H** (L) or **H** (L) with **L** (H), respectively, according to the specific condition.

---

<sup>7</sup> In the debriefing questionnaire, we asked to self-report about the number of experiments in which subjects had already participated in the past. The correlation between this self-reported measure and the actual record provided by ORSEE is 0.89. This shows how subjects are quite aware about their own individual level of lab-sophistication.

<sup>8</sup> In our framework, the within-subject design entailed a substantial boost in terms of statistical power [37].

<sup>9</sup> The randomization of C2 and C3 was important to enhance the robustness of our within-subject design [37] and the minimization of potential cognitive experimental demand effect [38].

**Table 1.** Experimental Design

Treatments	Matching Rule	Information
C1	Random H vs H	No information
C2	& L vs L	Information
C3	H vs L & L vs H	Information

We considered the following four games: Dictator Game, Ultimatum Game, Trust Game and Prisoner’s Dilemma Game<sup>10</sup>. In the first three asymmetric games, all the subjects played both role A (dictator/proposer/trustor) and role B (receiver/responder/trustee) in a strategy-method fashion, and subsequent stages were not announced in advance. As in [34], the games are played in a fixed sequence.<sup>11</sup> At the end of each experimental session, only one experimental condition, one game, one decision and one role for each player (when considering the dictator game, ultimatum game and trust game) were randomly selected and matched with their counterpart. Both players were paid in cash accordingly.

All subjects received 2.50 Euros as a show-up fee and got an average experimental payment of 7.50 Euros for a 45-minute lab session, including post-experimental surveys and debriefing. The exchange rate between ECU and Euros was 6 ECU=1 Euro. Six experimental sessions were conducted at the University of Cologne.<sup>12</sup> The experimental protocol was implemented using the Bonn Experiment System [39]. No feedback or results were received by participants before the end of the session.

#### 4. Testable Hypotheses

According to [22], the subjects’ behaviour changes over time as they get used to the experimental setting. This fact represents an issue both in terms of *intra-session* learning and *inter-sessions* familiarity and experience accumulation. As far as it regards intra-session learning, Binmore and Shaked [40] observe how the fact of getting used to the experimental setting leads the subjects to converge to behavioural patterns that are closer to the ‘homo economicus’ ones. Instead, Ding and Schotter [41] demonstrate how behaviour evolves via social learning and may diverge dramatically from that envisioned by the designer: experience with an incentive-compatible mechanism may not foster truthful revelation if that experience is achieved via social learning. Nevertheless, it is not clear

<sup>10</sup> The tasks are provided in the Appendix A.

<sup>11</sup> The logical sequence “Dictator Game → Ultimatum Game → Trust Game”, moving from the baseline case (DG) to the more complicated (TG) interaction, is implemented in order to favour the comprehension of the games and to avoid confusion. The Prisoner’s Dilemma game is placed at the end of the sequence in order to reduce the priming effect and because of the different nature of its dynamics.

<sup>12</sup> The instructions are provided in the Appendix B.

whether this is also the case for inter-sessions sophistication. Matthey and Regner [31] and Benndorf et al. [34] follow the same line in devising their hypothesis. From the conjecture that lab-sophisticated players are more inclined to play equilibrium strategies than lab-unsophisticated ones, and given our focus on fairness and reciprocity, we can derive the following general hypotheses:

**General Hypotheses:** Highly lab-sophisticated are more likely to play the payoff-maximizing strategies than low lab-sophisticated ones. That means that we should observe the following for high lab-sophisticated players compared to low lab-sophisticated players<sup>13</sup>: a) a smaller offer in the Dictator Game; b) a smaller average offer and a lower minimum acceptable offer (MAO) in the Ultimatum Games; c) a smaller average trust rate and a smaller average level of trustworthiness in the Trust Game; and, finally, d) a higher defection rate in the Prisoner’s Dilemma Game.

In testing these general hypotheses, we look separately at two possible channels of influence: sophistication and information. The “pure effect of lab-sophistication” is considered by comparing the behaviour of subjects with high and low lab-sophistication when they have no explicit information about their own and their counterpart’s level of sophistication. The “pure effect of information”, on the contrary, is considered by looking at changes in behaviour, given the same or different level of sophistication, when information is provided.

Given our experimental design, the analysis is structured as follows:

The “pure effect of lab-sophistication” is detected by comparing H-types and L-types in **C1** condition, where each participant is asked to make her decisions without receiving any information about the level of laboratory sophistication of the counterpart;

The “pure effect of information” becomes apparent when we compare the behaviour of one subject (e.g. H-type) when she is not informed about the type of the opponent (**C1**) vs. the behaviour of the very same subject when matched with an opponent of the same type (H-type in **C2** condition), or when she is matched with a different type (L-type in **C3** condition).

## 5. Results

In Table 2, we summarize the characteristics of the participants. We recruited 134 subjects (77 females and 57 males), with an average age of 25 and, balanced for the level of laboratory sophistication: 67 H-type subjects (min. 15, max. 86 previous participations in experiments, avg. 31) and 67 L-type subjects (min. 1, max. 5 previous participations in experiments, avg. 3). Our conventional sample size was set before any data analysis. Descriptive statistics show that, apart from the degree of laboratory sophistication, the two pools are fairly homogeneous.

---

<sup>13</sup> The following specific predictions are implied by the assumptions of selfish preferences and their being common knowledge.

**Table 2.** Balance of the Two Experimental Groups.

Characteristics	H types	L types	delta: $\Delta$ (H-L)
	(n = 67)	(n = 67)	<i>P-value</i>
	(n = 67)	(n = 67)	<i>P-value</i>
Female, #	37	40	0.60
(%)	(55.2 %)	(59.7 %)	
Age, mean	25.7	24.1	0.14
(min-max)	(19-60)	(18-65)	
Behavioural Economics classes, #	12	11	0.82
(%)	(18 %)	(16.4 %)	
Games Theory classes, #	23	18	0.35
(%)	(34.4 %)	(26.9 %)	

### 5.1. Lab-sophistication Effect

In the baseline condition (**C1**), players have no information about the lab-sophistication level of the counterpart. This experimental condition allows us to address our first question about the hypothesis of higher selfishness in H-type subjects with respect to L-type subjects.

Table 3 reports plain average outcomes by lab-sophistication level for the different games. In all the games, average outcomes generated by H-types and L-types look quite aligned and no systematic differences can be detected. The battery of between-subject non-parametric tests systematically fails in rejecting the null hypotheses of equality of means by level of sophistication (see Figures 1.a-f).

**Table 3.** Summary and Results (all games by sophistication levels – C1).

Games	Subjects type	Mean	Std. Dev	Median	Min	Max	MWU – Z	p-value
Dictator Game	H-type	<b>27.5</b>	22.6	30	0	80	0.815	0.415
	L-type	<b>31.1</b>	20.6	40	0	60		
Ultimatum Game Proposer	H-type	<b>39.6</b>	10.4	40	10	60	- 0.755	0.450
	L-type	<b>37.9</b>	10.4	40	0	100		
Ultimatum Game MAO Responder	H-type	<b>29.3</b>	13.9	30	0	50	-1.022	0.307
	L-type	<b>26.6</b>	14.7	30	0	50		
Trust Game Trustor	H-type	<b>23.6</b>	16.2	20	0	50	- 0.946	0.344
	L-type	<b>21.3</b>	16.6	20	0	50		
Trust Game Trustee <i>Individual average return</i>	H-type	<b>21.7</b>	14.0	26	0	40	1.072	0.283
	L-type	<b>24.4</b>	13.4	29	0	45		

Notes: Columns (8) and (9) are the results of a Wilcoxon-Mann-Whitney Nonparametric test.

Games	Subjects / Type	Defection Share	Std. Dev	Median	Min	Max	X <sup>2</sup>	p-value
Prisoner's Dilemma	H-type	<b>0.57</b>	0.5	1	0	1	0.496	0.481
	L-type	<b>0.63</b>	0.49	1	0	1		

Notes: Columns (8) and (9) are the results of Chi-square test, Pearson chi2(1).

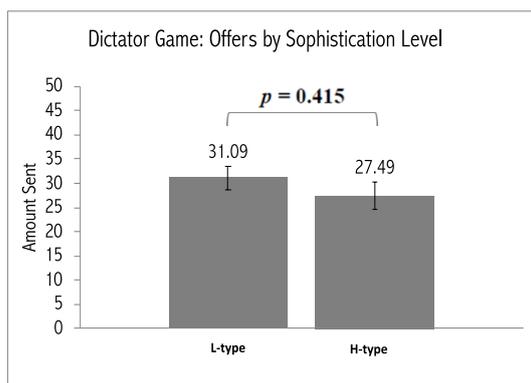


Figure 1.a: Dictator Game

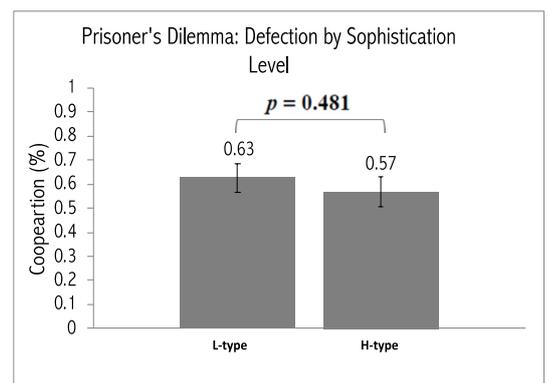


Figure 1.b: Prisoner's Dilemma

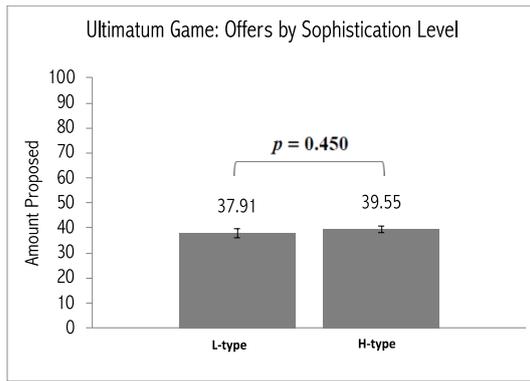


Figure 1.c: Ultimatum Game (Proposer)

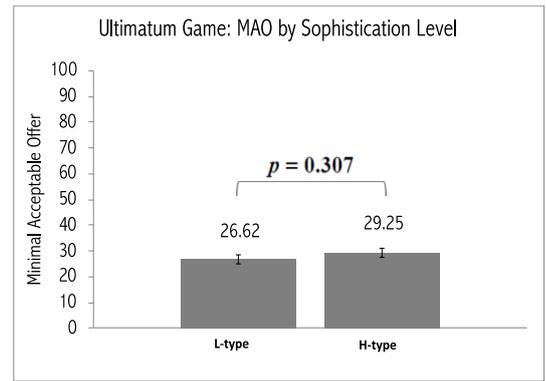


Figure 1.d: Ultimatum Game (Responder)

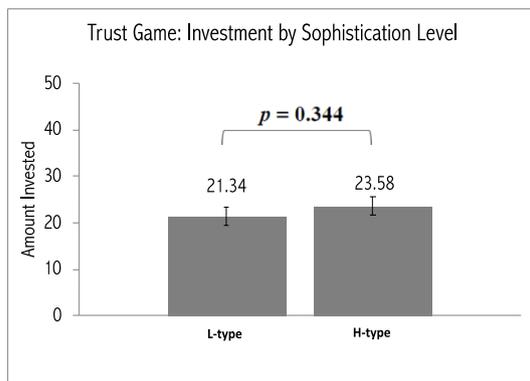


Figure 1.e: Trust Game (Trustor)

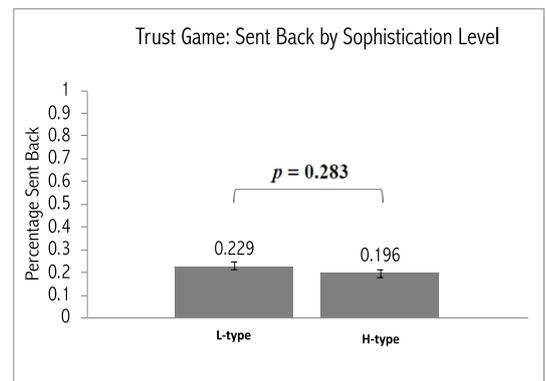


Figure 1.f: Trust Game (Trustee)

**Figure 1. a-f.** Choices in the Four Games by Sophistication Levels.

Table 4 goes one step further. It reports OLS regressions, allowing for a more comprehensive parametric analysis. In this second analytical assessment, all the outcomes under C1 are regressed – in a between-subjects fashion – on a dummy variable, named **H-dummy**, that identifies subjects with a high level of lab-sophistication. This coefficient informs us about the different behaviour of H-type subjects compared to L-types ones captured in the constant term.<sup>14</sup> We further take into account a battery of covariates including risk attitude, exposition to experimental/behavioural economics or game theory classes, gender, age, as well as more conventional controls, like monthly budget<sup>15</sup> and nationality.<sup>16</sup>

<sup>14</sup> While the size of the estimates associated with L-type (constant terms) and H-dummy coefficients does not quite match the absolute level of the corresponding averages reported in Table 4, due to partial effects captured by the control variables, from a qualitative perspective plain averages and parametric estimates are compatible.

<sup>15</sup> Dummy variables for different budget ranges: i) 0-300 €; ii) 301-1,000 €; iii) 1,501-2,000 €; iv) more than 2,001 €.

<sup>16</sup> Dummy variable =1 if non-German.

**Table 4.** Sophistication Effect (parametric analysis).

	(1)	(2)	(3)	(4)	(5)	(6)
outcome	DG	UG_roleA	MAO_UG_roleB	TG_roleA	TG_mean_roleB	PD
<b>H-dummy</b>	-3.789 (3.886)	2.041 (2.302)	3.378 (2.543)	3.145 (2.904)	-3.485 (2.523)	-0.078 (0.090)
risk_attitude	-1.123** (0.509)	-0.755** (0.301)	0.442 (0.335)	0.264 (0.380)	-0.352 (0.330)	-0.011 (0.012)
exp_class	-1.785 (5.485)	1.142 (3.248)	6.347* (3.566)	-6.284 (4.099)	0.648 (3.561)	-0.166 (0.127)
game_class	-2.006 (4.603)	-1.337 (2.726)	1.065 (3.001)	1.507 (3.440)	0.547 (2.988)	0.064 (0.107)
male	-6.928* (3.943)	4.277* (2.335)	3.899 (2.600)	8.051*** (2.947)	1.049 (2.560)	0.053 (0.091)
age	-0.127 (0.367)	0.055 (0.217)	0.038 (0.239)	-0.489* (0.274)	0.338 (0.238)	-0.001 (0.009)
controls	yes	yes	yes	yes	yes	yes
constant (L-type)	41.359*** (10.431)	39.993*** (6.178)	16.881** (6.772)	30.437*** (7.796)	18.816*** (6.772)	0.887*** (0.242)
Obs.	134	134	134	134	134	134
R-squared	0.109	0.139	0.145	0.136	0.049	0.081

Significance levels: \*  $p < 0.10$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ . Standard deviations reported in parenthesis.

In Column (1), dictator-game allocations are analyzed. The coefficient of H-dummy is not statistically significant at any conventional level. High-type dictators do not implement more greedy allocations compared to Low-type dictators. The allocation is marginally lower for males and partially affected by the individual degree of risk aversion.

Column (2) addresses ultimatum-game offers (role A). H-type proposers turn out to be neither more generous nor more selfish than the L-type ones. As for the dictators' allocations, ultimatum proposals are marginally affected by risk attitude, and at the same time marginally larger for male subjects. As far as it concerns the minimal acceptable offer (MAO) for ultimatum-game receivers (role B), the same tendency is confirmed (see Column [3]).

Column (4) focuses on trust-game offers (role A). Also in this case, H-type subjects show a behavioural pattern that is comparable with the one exhibited by L-type subjects. Trust appears to be significantly higher in males and marginally negatively affected by age.

The same result is confirmed when the behaviour of trustees (role B) is addressed. In Column (5), the outcome is represented by the mean amount returned by responders and, it is computed at

individual level, averaging the trustees' multiple responses under the different trust profiles (strategy method). As for the previous outcomes, it is not possible to detect any significant difference in the estimated return rates of the two different populations; this indicates the absence of any differential effect due to a higher level of lab-sophistication.

Column (6) analyses prisoner's dilemma strategies, implementing a linear probability model. L-type subjects defect in the majority of the cases (constant term), and the H-dummy variable does not prove statistically significant difference between the two subjects' population.

The general behavioural pattern that emerges from this first analysis proves to be quite consistent across the six different archetypal interactions: H-type subjects do not prove to be systematically more selfish than L-type subjects. A high individual level of experimental lab-sophistication per se does not systematically push subjects towards more selfish choices.

## 5.2. Information Effect

Exploiting within-subject outcome variations, we now focus on how individual behaviour changes when information about the type of the opponent is revealed. In order to do that, we contrast the outcome observed under the two further experimental manipulations, **C2** and **C3**, against the baseline outcome observed under **C1**, where the sophistication level of the opponent is unknown.

Table 5 reports GLS regressions with random effects at subject level, as we are now addressing within-subject outcome variations (see [42]).

L-types and H-types subjects are analysed separately. All the outcomes are regressed on the two "condition" dummies, **C2** – *same* and **C3** – *different*, which identify the opponent's level of laboratory sophistication (L or H types). Under **C2** condition, subjects are matched with players of the same type (H- types vs. H-types/ L-types vs. L-types), while under **C3** condition the matching is heterogeneous (H-types vs. L-types/ L-types vs. H-types). We further control for a conventional battery of covariates, including risk attitude, exposition to experimental/behavioural game-theory classes, gender, age, and controls like monthly budget and nationality.

The first two columns address **dictator-game** allocations. Column (1) addresses L-type dictator allocations. In this case, the **C2** – *same* dummy informs us about the change in the L-type dictator allocation when she is informed of being matched with a L-type recipient. The coefficient associated with **C2** – *same* turned out to be small in its magnitude and not statistically significant at any conventional level. This means that, on average, the behaviour of L-type dictators does not change when the recipient's low level of sophistication is disclosed. Following the same line of reasoning, the **C3** – *different* dummy informs us about the change in the L-type dictator's allocation when the dictator is informed of being matched with a H-type recipient. The coefficient for **C3** – *different* turned out to be negative, but rather small in its magnitude (-4 points) and, only marginally statistically significant ( $p < 0.10$ ). This means that, on average, L-type dictators tend to give a bit less when the recipient's high level of lab-sophistication is disclosed. Column (2) reports the very same assessment for H-type dictators. Both in the case in which it is revealed she is interacting with a H-type and the case in which it is revealed she is dealing with an L-type, no statistically significant differences in dictator's allocations are detected.

Columns (3) and (4) address **ultimatum-game** allocations (**role A**) for L-type and H-type subjects, respectively. In column (3) for L-type proposers, both coefficients of **C2** – *same* and **C3** – *different* dummies prove to be negative, only marginally statistically significant at the 10% level and similar in their own sizes (-3 points, Wald test  $p > 0.80$ ). This common departure from the baseline

(C1 condition) is irrespective of the type of the opponent. In column (4) for H-type proposers, both coefficients of *C2 – same* and *C3 – different* dummies are negative and statistically significant at the 5% and 1% level, respectively. H-type proposers give less when information is provided, irrespectively of the type of opponent; while the *C2 – same* dummy coefficient brings a negative effect by -2 points, and *C3 – different* by -4 points, the Wald test fails to reject the null hypothesis of equality ( $p > 0.10$ ).

Columns (5) and (6) analyze the minimal-acceptable-offer (MAO) for L-type and H-type **ultimatum-game** recipients (**role B**), respectively. On the one hand, L-type recipients matched with L-type proposers are willing to accept slightly smaller offers (-2.8 points) than in the baseline condition. The negative effect associated with the *C2 – same* dummy coefficient is significant at the 5% level. On the other hand, when an L-type recipient is interacting with a H-type proposer (captured by *C3 – different* dummy coefficient), no differential effect can be detected with regard to the baseline case (column [5]). The same analysis for H-type ultimatum-game recipients is described in column [6]. Both the estimates for the variables dummies *C2 – same* and *C3 – different* turn out to be negligible in their magnitudes and no significant statistical differences can be observed compared to the benchmark condition.

Columns (7) and (8) focus on **trust-game** offers (**role A**) of L-type and H-type subjects, respectively. The L-type trustor, in column (9), does not seem to be sensitive to the level of sophistication of the trustee with whom she is interacting. Both coefficients of *C2 – same* and *C3 – different* dummies are small in size and not statistically significant at any conventional level. As a side remark, we detect the negative effect associated with the dummy variable signalling the previous exposition to experimental-economics classes. In column [8], H-type trustors seem to trust less when they face L-type trustees. The coefficient of *C3 – different* dummy, in fact, is negative and statistically significant at the 1% level. The negative effect has an estimated magnitude of -4 points sent to L-type trustees. Males turn out to be significantly more trustful, while age brings a marginal negative.

Trustworthy responses by L-type and H-type trustees (**role B**) are reported in columns (9) and [10], respectively. While L-type responders apply the very same level of trustworthiness in the three different conditions (in column (9), the coefficients of *C2 – same* and *C3 – different* dummies are small and not statistically different from the constant term C1), H-type subjects marginally decrease their level of trustworthiness when the trustor's type is revealed. In column [10], the estimate for *C2 – same* dummy variable is equal to -1.3 points and -2 points for *C3 – different* dummy variable. In both cases, the negative effect turns out to be statistically significant at the 5% and 10% level, respectively.

Column (11) focuses on **prisoner's dilemma** strategies<sup>17</sup> for L-types. The defection strategy proves to be the most adopted one under the three different experimental conditions. No significant departure from the baseline can be observed, either under C2 or in the C3 condition. Column [12] looks at H-type subjects behaviour. Also, in this case, the defection strategy is the most frequently implemented option, with just a marginal variation. Under C3 condition, captured by *C3 – different* dummy coefficient, H-types tend to be marginally ( $p < 0.10$ ) more cooperative (+7 percentage points) when matched with an L-type.

In sum, although no clear and consistent behavioural pattern emerges, we can say that the exogenous provision of information about the level of lab-sophistication of the counterpart seems to affect the H-type participants' choices mostly. In particular:

<sup>17</sup> The binary outcome assumes value 1 in case of defection, 0 in case of cooperation.

- in the ultimatum game, for the H-type proposers, both coefficients for C2 (same type match: H-type vs. H-type) and C3 (different type match: H-type vs. L-type) prove to be negative and statistically significant at the 5% and 1% level, respectively. H-type proposers give less when information is provided, irrespectively of the type of opponent;
- in the trust game, H-type trustors invest less in L-type trustees, and H-type trustees send back less when the trustor's type is revealed, irrespectively of her type;
- finally, in the prisoner's dilemma game, H-types tend to be more cooperative when they know they are interacting with an L-type.

As far as it concerns the L-type participants, we only find that L-type dictators tend to give a bit less to H-type recipients ( $p < 0.10$ ); in the Ultimatum Game, as far as the receivers' behaviour is concerned, L-types are willing to accept lower offers when they know that they come from L-type proposers ( $p < 0.05$ ).

**Table 5.** Information Effect (parametric analysis).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Outcome	DG_	DG_	UG_	UG_	UG_MAO_	UG_MAO_	TG_	TG_	TG_mean_	TG_mean_	PD_	PD_
	L-type	H-type	roleA_	roleA_	roleB_	roleB_	roleA_	roleA_	roleB_	roleB_	L-type	H-type
			L-type	H-type	L-type	H-type	L-type	H-type	L-type	H-type		
<i>C2 – same</i>	-1.194	-1.522	-3.134*	-2.388**	-2.894**	-0.299	0.001	-0.448	-0.567	-1.352**	0.030	0.001
	(2.46)	(1.95)	(1.85)	(1.09)	(1.34)	(0.85)	(1.38)	(1.38)	(1.13)	(0.66)	(0.06)	(0.05)
<i>C3 – different</i>	-4.075*	-2.343	-3.433*	-4.179***	-0.313	0.299	-0.896	-4.030***	-1.552	-2.101***	-0.001	0.075*
	(2.46)	(1.95)	(1.85)	(1.09)	(1.33)	(0.85)	(1.38)	(1.38)	(1.13)	(0.66)	(0.06)	(0.05)
risk_attitude	-0.460	-1.212	-0.048	-0.484	0.788*	0.150	0.613	0.346	0.340	-0.897*	-0.026*	0.001
	(0.64)	(0.76)	(0.48)	(0.42)	(0.48)	(0.49)	(0.53)	(0.56)	(0.50)	(0.53)	(0.01)	(0.02)
exp_class	-4.505	-0.827	-7.905	1.795	4.898	4.397	15.470***	0.024	-8.048	5.368	0.055	-0.249
	(6.76)	(8.01)	(5.02)	(4.47)	(5.01)	(5.15)	(5.63)	(5.97)	(5.26)	(5.57)	(0.15)	(0.18)
game_class	-5.333	-0.455	-0.334	-2.302	-0.710	0.915	1.857	0.084	0.280	-3.332	0.021	0.027
	(5.75)	(6.19)	(4.27)	(3.46)	(4.26)	(3.98)	(4.79)	(4.61)	(4.48)	(4.31)	(0.12)	(0.14)
male	-1.073	-4.968	2.799	4.741	1.388	4.386	1.982	9.806**	3.964	0.164	-0.076	0.068
	(5.03)	(5.27)	(3.74)	(2.94)	(3.74)	(3.39)	(4.20)	(3.93)	(3.92)	(3.66)	(0.11)	(0.12)
age	-0.299	-0.020	0.220	-0.070	0.163	-0.001	0.123	-0.591*	-0.091	0.050	0.006	0.001
	(0.80)	(0.43)	(0.60)	(0.24)	(0.60)	(0.28)	(0.67)	(0.32)	(0.63)	(0.30)	(0.02)	(0.01)
controls	yes	Yes	yes	yes	Yes	yes	yes	Yes	yes	yes	yes	yes
random effects	yes	Yes	yes	yes	Yes	yes	yes	Yes	yes	yes	yes	yes
constant (C1)	49.796***	28.674**	38.602***	42.178***	10.527	27.997***	22.984	30.950***	29.238**	21.461**	0.627	0.694**
	(19.07)	(14.44)	(14.18)	(8.06)	(14.12)	(9.26)	(15.88)	(10.76)	(14.82)	(10.02)	(0.41)	(0.32)
Obs.	201	201	201	201	201	201	201	201	201	201	201	201
R-squared	0.097	0.094	0.098	0.127	0.175	0.134	0.165	0.158	0.07	0.153	0.073	0.094

Significance levels: \*  $p < 0.10$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ . Standard deviations reported in parenthesis.

## 6. Conclusions

The generalizability of conclusions drawn from lab experiments is still a debated issue in economics. It is, of course, a multifaceted problem that refers to many dimensions of the experimental practices and methods: the artificiality of the situations considered in the lab, the small size of the incentives, and the lack of representativeness of the experimental subjects are only a few of the problematic elements. In particular when we consider the reliability of the conclusions drawn from experiments involving a convenience pool of students, we should also ask whether the repeated participation into different experiments by these subjects might have a lasting effect on their behavioural tendencies in the lab, even regardless of the nature of the tasks they were involved in the previous experiments. Were this to be true, in fact, lab-sophisticated subjects would constitute an even less representative pool whose behaviour patterns could not be reliably generalized. In this paper, we precisely addressed this point. By design, we investigated whether having repeatedly taken part in previous experiments consistently modifies the behaviour of individuals, when compared to sophisticated (H-type) or unsophisticated (L-type) participants, in a set of widely used games: dictator game, ultimatum game, trust game, and prisoner's dilemma game. We considered a between-subjects design to compare the behaviour of sophisticated (H-type) or unsophisticated (L-type) subjects in the four games, and we exploited a within-subjects design to explore the extent to which knowledge about the level of laboratory sophistication of the counterpart affects players' decisions.

While sophistication *per se* seems not to have any significant effect on participants' pro-social and cooperative choices, our data suggest how knowing the level of sophistication of the co-players may systematically alter the behaviour of H-type participants, especially in the ultimatum game and in the trust game. We do not advance any theoretical explanation for this phenomenon; we only stress that fact that if such information can be inferred by the participants, due to the characteristics of the recruitment mechanisms or the management of the experimental sessions, this could have a negative effect on the reliability of the laboratory data. This result should induce experimenters to take into account a further methodological prescription and to pay attention to neutralize all contextual and procedural details from which each individual participant could draw information about the level of sophistication of all the other participants in order to minimize the risk of measuring systematically distorted behavioural.

**Appendix:** A and B, online supplementary materials – link: [tinyurl.com/k0nzupwb](http://tinyurl.com/k0nzupwb)

**Acknowledgments:** We thank Enrique Fatas for his encouragement and stimulus in the early stages of this project; the staff of CODEBE and CLER Lab of the University of Cologne for research assistantship; Valerio Capraro, Miloš Fišar, Francesco Guala, Andrea Isoni, Ro'i Zultan and the participants at the "Methods" parallel session of 2015 ESA European Conference (Heidelberg), for their comments and suggestions. We are grateful to the Editors (Riccardo Ghidoni - Ulrich Berger) and three anonymous referees for their comments and suggestions that helped us to improve the manuscript.

## References

1. Sugden, R. (2005). Experiment, theory, world: A symposium on the role of experiments in economics. *Journal of Economic Methodology*, 12(2), 177-184.
2. Binmore, K., (1999), Why Experiment in Economics? *The Economic Journal*, 109, F16-F24.
3. Guala, F., & Mittone, L. (2005). Experiments in economics: External validity and the robustness of phenomena. *Journal of Economic Methodology*, 12(4), 495-515.
4. Schram, A. (2005). Artificiality: The tension between internal and external validity in economic experiments. *Journal of Economic Methodology*, 12(2), 225-237.
5. Cooper, D. J., Kagel, J. H., Lo, W., and Gu, Q. L. (1999). Gaming against Managers in Incentive Systems: Experimental Results with Chinese Students and Chinese Managers. *The American Economic Review*, 89(4), 781-804.
6. Peterson, R. A. (2001). On the Use of College Students in Social Science Research: Insights from a Second-Order Meta-Analysis. *Journal of Consumer Research*, 28, 450-461.
7. Fehr, E. and List J. A. (2004). The Hidden Costs and Returns of Incentives. Trust and Trustworthiness among CEOs. *Journal of the European Economic Association*, 2(5), 743-771.
8. Carpenter, J., Burks, S., and Verhoogen, E. (2005). Comparing students to workers: The effects of social framing on behavior in distribution games. *Field Experiments in Economics (Research in Experimental Economics)*, Carpenter, J., Harrison, G., List, J. (editors), JAI/Elsevier, Greenwich, CT, and London, pp. 261–290.
9. Bellemare, C. and Kroger, S. (2007). On representative social capital. *European Economic Review*, 51, 183–202.
10. Danielson, A. J. and Holm, H. J. (2007). Do You Trust Your Brethren? Eliciting Trust Attitudes and Trust Behavior in a Tanzanian Congregation. *Journal of Economic Behavior and Organization*, 62(2), 255-271.
11. Alatas, V., Cameron, L., Chaudhuri, A., Erkal, N., and Gangadharan, L. (2009). Subject pool effects in a corruption experiment: A comparison of Indonesian public servants and Indonesian students. *Experimental Economics*, 12(1), 113-132.
12. Belot, M., Duch, R. M., and Miller, L. M. (2010). Who Should be Called to the Lab? A comprehensive comparison of students and non-students in classic experimental games. ([http://cess-wb.nuff.ox.ac.uk/documents/DP2010/CESS\\_DP2010\\_001.pdf](http://cess-wb.nuff.ox.ac.uk/documents/DP2010/CESS_DP2010_001.pdf))
13. Cappelen, A. W., Nygaard, K., Sørensen, E. Ø., and Tungodden, B. (2015). Social Preferences in the Lab: A Comparison of Students and a Representative Population. *The Scandinavian Journal of Economics*, 117(4), 1306-1326.
14. Anderson, J., Burks, S. V., Carpenter, J., Götte, L., Maurer, K., Nosenzo, D., Potter, R., Rocha, K., and Rustichini, A. (2013). Self-selection and variations in the laboratory measurement of other-regarding preferences across subject pools: Evidence from one college student and two adult samples. *Experimental Economics*, 16 (2), 170-189.
15. Fréchette, G. R. (2015). Laboratory Experiments: Professionals versus Students. In *Handbook of Experimental Economic Methodology*, Guillaume R. Fréchette and Andrew Schotter (editors), Oxford University Press, February, 360-390.
16. Henrich, J., Heine, S. J., and Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2-3), 61-83.
17. Eckel, C. and Grossman, P. J. (2000). Volunteers and pseudo-volunteers: The effect of recruitment method in dictator experiments. *Experimental Economics*, 3, 101-120.

18. Falk, A., Meier, S., and Zehnder, C. (2010). Do Lab Experiments Misrepresent Social Preferences? The case of self-selected student samples. *Journal of the European Economic Association*, 11(4), 839-852.
19. Cleave, B. L., Nikiforakis, N., & Slonim, R. (2013). Is there selection bias in laboratory experiments? The case of social and risk preferences. *Experimental Economics*, 16(3), 372- 382.
20. Exadaktylos, F., Espín, A. M., and Brañas-Garza, P. (2013). Experimental subjects are not different. *Scientific Reports*, 3, 1213.
21. Guillen, P., & Veszteg, R. F. (2012). On “lab rats”. *The Journal of Socio-Economics*, 41(5), 714-720.
22. Friedman, D., and Cassar, A. (2004) *Economics Lab: An intensive course in experimental economics*, Routledge, 2004.
23. VandenBos, G. R. (Ed.). (2007). *APA Dictionary of Psychology: ‘Subject Sophistication’*. American Psychological Association.
24. Kruglanski, A. W. (1975). The human subject in the psychology experiment: Fact and artifact. In *Advances in experimental social psychology* (Vol. 8, pp. 101-147). Academic Press.
25. Greiner, B., & Stephanides, M. (2019). Subject pools and recruitment. In: *Handbook of Research Methods and Applications in Experimental Economics*. Edward Elgar Publishing.
26. Harrison, G. W., McKee, M., and Rutström, E. E. (1987). Experimental evaluations of institutions of monopoly restraints. In: *Advances in Behavioral Economics*, Green, L., Kagel, J. H. (editors), Ablex Press, Norwood, N.J., 54-94.
27. Benson, B. L., and Faminow, M. D., (1988). The impact of experience on prices and profits in experimental duopoly markets. *Journal of Economic Behavior and Organization*, 9(4), 345-365.
28. Marwell, G., and Ames, R. E., (1980). Experiments on the provision of public goods. II. Provision points, stakes, experience and the free rider problem. *American Journal of Sociology*, 85(4), 926-937.
29. Isaac, M., Walker J. M., and Thomas S. H. (1984). Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations. *Public Choice*, 43, 113-149.
30. Bolton, G. E. (1991). A comparative model of bargaining: theory and evidence. *American Economic Review*, 81(5), 1096-1136.
31. Matthey, A., and Regner, T. (2013). On the independence of history: Experience spill-overs between experiments. *Theory and Decision*, 75, 403-419.
32. Capraro, V. and Cococcioni, G. (2015). Social setting, intuition, and experience in lab experiments interact to shape cooperative decision-making. *Proceedings of the Royal Society B: Biological Sciences*, 282(1811).
33. Xue, L., Sitzia, S. and Turocy, T. L. (2015). Mathematics self-confidence and the “prepayment effect” in riskless choices. <https://www.uea.ac.uk/documents/166500/0/CBESS+15-20.pdf/31995582-7f0a-4ab6-bf72-7b53a9e6cab5>
34. Benndorf, V., Moellers, H. C., Normann, H. (2017). Experienced vs. inexperienced participants in the lab: do they behave differently? *Journal of the Economic Science Association*, 3(1), 12-25.
35. Conte, A., Levati, M. V., and Montanari, N. (2019). Experience in Public Goods Experiments. *Theory and Decision*, 86, 65–93.
36. Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1), 114-125.

- 
37. Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, 81(1), 1-8.
  38. Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, 13(1), 75-98.
  39. Seithe, M. (2012). Introducing the Bonn Experiment System [http://boxs.uni-bonn.de/boxs\\_seithe.pdf](http://boxs.uni-bonn.de/boxs_seithe.pdf).
  40. Binmore, K., and Shaked, A., (2010), Experimental economics: Where next? *Journal of Economic Behavior & Organization*, 73(2010), 87–100.
  41. Ding, T., & Schotter, A. (2019). Learning and mechanism design: An experimental test of school matching mechanisms with intergenerational advice. *The Economic Journal*, 129(623), 2779-2804.
  42. Hennig-Schmidt, H., Irlenbusch, B., Rilke, R. M., & Walkowitz, G. (2018). Asymmetric outside options in ultimatum bargaining: A systematic analysis. *International Journal of Game Theory*, 47(1), 301-329.