Experimental Design of iLEE1

August 16, 2012

Abstract
This document describes the design of iLEE1, the first wave of the Internet Laboratory for Experimental Economics project at the University of Copenhagen. The project is directed by Jean-Robert Tyran and funded by the Carlsberg Foundation. Nikos Korfiatis was responsible for coding the web interface and Erik Wengström was in charge of implementing all other operational aspects of iLEE1. We thank Toke Fosgaard, Lars Gårn Hansen, Eva Gregersen, Ditte Morup, Louise Skouby and Anja Skadkær Møller and Statistics Denmark for invaluable support and assistance.

For specific questions, please contact:
Erik Wengström
Department of Economics
Lund University
P.O. Box 7082
S-220 07 Lund, Sweden
Email: erik.wengstrom@nek.lu.se

For general questions, please contact:
Jean-Robert Tyran
Department of Economics
University of Vienna
Hohenstauengasse 9
A-1010 Vienna, Austria
Email: jean-robert.tyran@univie.ac.at
1 Introduction

This document describes the design of a large-scale public good game carried out over the internet in Denmark in May 2008.

2 Recruitment of subjects

The participants were recruited in the following way:

1. In November 2007, we asked the interview department at Statistics Denmark to randomly select 40,000 individuals from the Danish population.

2. We provided the text for the invitation letters and Statistics Denmark printed the respondent’s name and address together with a personal identification number on each of the letters. See Appendix A for a picture of one of these invitation letters and the envelope used to invite individuals to participate in the experiment. Letters together with window envelopes were sent back to us. We hired students to pack the letters and sort them randomly into boxes of 150-250 letters in each. The scheme according to which the letters were divided into boxes was constructed using the random number generator in Microsoft Excel.

3. A first set of 2503 invitation letters were sent out on May 15. The letter invited subjects to log on to our webpage, www.econ.ku.dk/ilee using the personal identification number printed in the letter. A second wave consisting of 19,524 letters were sent out the following week on May 20. Participants could log on to our web page and participate in our experiment within seven days after receiving the invitation letter. Out of the 22,027 invitation letters, 4,290 subjects logged on to our web page and 2,291 subjects completed the experiment.

3 Register Data at Statistics Denmark

After the experiment was conducted the data was sent to Denmark Statistics. They used the anonymous log in codes to match the experimental data with a battery of register data. The data is stored in an anonymous format at a server at Denmark’s Statistics.

4 Experimental design

4.1 Overview

Figures 1 and 2 illustrate the screen-by-screen flow of our experiment. More detailed information about each of these steps follows in Section 4.3. Translated screenshots of the full experiment are available in appendix B.

Note that citizens who have claimed research protection are not included in the population from which our sample of 40,000 was drawn.
In short, the participants were invited to log on to our web page twice, once during the period in which the experiment was open and once during a feedback period after the experiment was closed. The first time they logged in they participated in the public goods game, the risk and loss task and completed the other questionnaires and tests described in Figure 1. After the experiment closed, participants were matched together in groups for the public good game and payments were calculated. Participants logged on to our web page again to see the results for their group and provide us with the banking details necessary for distributing the payments. The contents of this feedback part are depicted in Figure 2.
Figure 1: Part 1 – The experiment

1. Introduction
   (a) Login screen
   (b) Welcome screen
   (c) Preliminary background questions

2. The public good game
   (a) Instructions stage 1 (pure one-shot)
   (b) Control questions
   (c) Stage 1 - Unconditional choice
   (d) Belief elicitation
   (e) Instructions stage 2 (strategy method)
   (f) Stage 2 - Unconditional choice
   (g) Stage 2 - Conditional choices
   (h) Confusion questions

3. Risk and loss preferences
   (a) Instructions
   (b) Risk aversion elicitation
   (c) Loss aversion elicitation

4. Background questions
   (a) Workplace
   (b) IT
   (c) Political
   (d) World Value Survey - trust, fairness and happiness

5. Personality and cognitive tests
   (a) Cognitive reflection test
   (b) Big Five personality test
   (c) Intelligence-Structure-Test 2000 R (Matrices)

6. End of first login stage
   (a) Control number and email request
   (b) Additional comments
1. **Introduction**

(a) Login screen
(b) Welcome back and control code screen
(c) Results screens
(d) Banking details
(e) Additional comments

### 4.2 Treatments

The experiment had six treatments that varied with respect to invitation letter type, incentives and framing.

There were three types of invitation letters, namely Standard, Support, and Support Gift. The Standard letter informs subjects that they can make money in the experiment, whereas the Support letters instead tell subjects that they will be contributing to scientific research. See Figure 3 in Appendix A for a picture of the invitation letter and Appendix B for translations of the text on the letters. The Support Gift letter also included a small gift (a foam puzzle with the iLEE and University of Copenhagen logos printed on it). See Figure 5 in Appendix A for a picture of this gift.

Another variation concerned the actual incentives paid out to the subjects. In the Paid treatments, subjects actually received payment irrespective of which invitation letter they received. In the Hypothetical treatments, subjects faced the same instructions throughout the experiment, but the welcome screen included a paragraph making clear that subjects would not actually receive payment and directing them to simply make their decisions as if they would be paid according to the instructions. Naturally only subjects receiving invitation letter Support and Support Gift participated in the Hypothetical treatment.

The third and last variation concerned the framing of the public good game part of the experiment, which was either a Give or a Take frame. Only the instructions for the public good game differed between treatments. See Section 4.3 for an explanation of the difference.

Table 1 breaks down the complete target subject pool into treatments.

<table>
<thead>
<tr>
<th>Letter Type</th>
<th>Letters sent out</th>
<th>Subjects logging in ( completing the experiment)</th>
<th>Paid</th>
<th>Give</th>
<th>Hypothetical Give</th>
<th>Paid</th>
<th>Take</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>18,027</td>
<td>2,027 (1,366)</td>
<td>-</td>
<td>1,080 (676)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Support</td>
<td>2,000</td>
<td>93 (47)</td>
<td>128 (68)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Support Gift</td>
<td>2,000</td>
<td>146 (85)</td>
<td>110 (49)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>22,027</td>
<td>2,266 (1,498)</td>
<td>238 (117)</td>
<td>1,080 (676)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: Figures refers to the number of subjects logging into the webpage. Numbers in parenthesis refers to number of subjects completing the experiment.*

### 4.3 Detailed design descriptions

This section describes each part of the experiment in detail.
4.3.1 Login screen

The first screen of the experiment that subjects were taken to when they entered the URL from the invitation letter was a simple login screen where subjects had to enter the personal identification code also found in the letter.

4.3.2 Welcome screen

Upon login, subjects saw a welcome screen providing information about the experiment. They were informed that their participation in the experiment would be valuable to research in economics and reminded of the importance that the person participating was the person named in the invitation letter. If the subject had been assigned to a Paid treatment, they were informed that they could earn money in the experiment (within the range of 8 to 510 DKK (∼1 to 68 EUR)) and that this is standard procedure in economic experiments. They were also cautioned that they had to complete the experiment to get their money. If the subject had been assigned to a Hypothetical treatment, they were instead instructed that although they would seemingly be earning money in the experiment, this would not be paid out to them, but that they should try to make decisions as if they were really earning the money. All subjects were then informed that the experiment would last approximately 50 minutes. Finally, they were reassured that they would be anonymous.

4.3.3 Preliminary background questions

The first screen after the welcome screen contained three preliminary background questions about subjects’ age, gender and highest obtained level of education. These three background questions were placed at the beginning of the experiment when minimal attrition had taken place such that we might validate the identities of as many subjects as possible. By comparing a subject’s stated demographic data with the (anonymous) data from Statistics Denmark for the individual to whom the invitation letter was addressed, we obtain a proxy control against the participating subject being the wrong individual.

An additional reason to ask these three particular background questions is that it was required for our collaboration with Dansk Psykologisk Forlag (DPF), a Danish non-profit foundation that develops, translates and publishes materials for psychologists. DPF permitted us to use copyrighted personality and IQ tests (see below) in exchange for getting the data for each test along with those three self-reported background variables.

After completing the preliminary background questions, subjects proceeded to the public good game part of the experiment.

4.3.4 The public good game

Subjects played two variants of the public good game. They first played a standard linear one-shot public good game involving one unconditional contribution choice. Afterwards they played a version using the strategy method which involves an unconditional choice as well as a series of conditional choices.

In both games, there were four members in each group, the endowment was 50 DKK (∼7 EUR), and the marginal per capita return was 0.5. In the Give frame subjects were asked to contribute between 0-50
DKK of the private endowment to a common pot. Everything in the pot was then doubled and shared equally between the four subjects in the group. In the Take frame 200 DKK were initially placed in the public pot and subjects were asked to withdraw between 0-50 DKK from the pot. Analogous to the Give frame, everything left in the pot was thereafter doubled and shared equally between the members of the group. The frame (Give or Take) that each subject encountered was also constant. There was no feedback during game play.

Subjects began by reading the instructions for the unconditional choice. In order to make the rules of the public good game easy to understand, the written instructions were complemented by a series of illustrations made by a professional illustrator. Subjects had access to several forms of help in understanding the instructions. Throughout the public good game part of the experiment, subjects could go back and read the instructions again at any time. In addition, from each screen subjects could access a screen-specific help screen which provided further guidance about what to do. Subjects also had access to a profit calculator where they could see for themselves how the earnings of the four members of the group depended on the members' contributions. Finally, all help screens included a telephone number and an email address through which subjects could obtain further assistance.

After viewing the instructions, subjects were required to correctly complete four control questions testing their ability to calculate payoffs in the game. Subjects were allowed as many tries as necessary but could not proceed without entering the correct answer to each question.

Subjects then made their first unconditional choice.

On the next screen, their beliefs about the average contribution of the other members of their group were elicited. The belief elicitation was incentivised using the quadratic scoring rule.

Subjects then read the instructions for the strategy method version of the public good game. The strategy method was adapted to the context of the public good game by Fischbacher, Gachter, and Fehr (2001). The idea behind the strategy method is to have subjects report the complete strategy of actions they would like to take in the event of each possible combination of actions that others could take.

After reading the instructions, subjects first had to make a second unconditional choice. This unconditional choice was necessary to determine the outcome of the game.

Subjects then had to fill out a conditional contribution table in which they had to decide how much they would like to contribute for each of the 11 average contribution levels of the other group members that are multiples of 5 (0, 5, 10... 45, 50). In this respect our design differs from Fischbacher et. al. (2001) where the endowment was 20 tokens and where all 21 possible integer average contribution levels were included in the conditional contribution table.

The outcome of this second public good game was determined as follows: One member of the group is randomly selected. For the other three subjects, the second unconditional choice counts as their contribution. The average of their choices is rounded to the nearest multiple of 5, and the contribution of the selected member is then determined by referencing the relevant row of his or her conditional contribution table.

Finally, subjects completed six incentivised confusion questions. The first three questions tested
their understanding of the dominant strategy, and the last three questions tested their understanding of the socially optimal strategy. Each correct answer earned the subject 5 DKK (≈0.7 EUR). These confusion questions concluded the public good game part of the experiment.

4.3.5 Risk and loss preferences

The next part of the experiment involved the elicitation of risk and loss preferences. The measures used were lottery tasks adapted from Tanaka, Camerer, and Nguyen (2009).

Subjects first read the instructions which directed them to make two series of choices between two games, Left or Right, of Heads or Tails, i.e. binary lotteries with equally likely outcomes. The payoffs are constructed with greater variance and increasing expected utility in the Right column. A risk (loss) neutral individual would choose to start by playing the games in the Left column and then switch to playing the games in the Right column as soon as the expected utility in that column started to outperform that of the Left column, whereas a risk (loss) averse individual would choose to linger in the Left column. The degrees of risk and loss aversion can therefore be defined as the total number of Left games chosen.

After reading the instructions, subjects proceeded to the 10 lottery choices pertaining to risk aversion and subsequently to the 7 choices pertaining to loss aversion.

Our design differs from Tanaka, Camerer, and Nguyen (2009) in four respects. First, the outcomes of the binary lottery games are not equally likely with respect to risk aversion in their design. We introduce 50%-50% gambles since we were not primarily interested in the issue of probability weighting. In addition, using 50%-50% gambles simplified the task for the subjects. Second, due to time restrictions we only used one table of choices to measure risk aversion whereas Tanaka, Camerer, and Nguyen (2009) use two tables. Third, we changed the payoffs involved in each game. Finally, Tanaka, Camerer, and Nguyen (2009) enforce monotonic switching by asking subjects only at which point, if any, they would switch between the two columns of games, whereas we ask subjects to make a choice between the games in each individual row in order to get some measure of the noisiness in subjects’ choices.

The measures were incentivised as follows: One row out of the total of 17 rows in both games was randomly selected, and the chosen game for that row was played. If the outcome was one in which the subject incurred a loss, then the amount in question was subtracted from their total earnings in the experiment.

4.3.6 Background questions

After the measures of risk and loss aversion came a second set of background questions. We used the questions about trust, fairness, happiness and political values questions from the WVS. We also asked questions about computer usage, interaction in online social networks and workplace conditions. See the instructions for details.
4.3.7 Cognitive Reflection Test

The Cognitive Reflection Test (CRT), is a short three-question test aimed at capturing a specific dimension of cognitive ability. According to Frederick (2005) who introduced the test it captures individuals’ ability or disposition to reflect on a question and resist reporting the first response that comes to mind.

The test is composed of the following three questions:

1. A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost? .... cents

2. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? .... minutes

3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? .... days

The three CRT problems generate the incorrect "intuitive" answers (10, 100, and 24), but the correct ones are (5, 5, and 47). Test scores are calculated as the number of correct answers, i.e. 0, 1, 2, or 3. Frederick (2005) compares the CRT scores with other measures of cognitive ability such as the Scholastic Achievement Test (SAT) and the American College Test (ACT). These are the two most common college entrance examinations in the U.S and are also described as measures of academic achievement. The author shows, that all cognitive measures correlate positively and significantly with each other, suggesting that all tests likely reflect common factors, but may also reflect distinct characteristics. For example, though the CRT is intended to measure cognitive reflection, performance is surely aided by reading comprehension and mathematical skills (which the ACT and SAT also measure).

4.3.8 Big-five personality test

The Big Five Inventory (BFI) is the most prominent measurement system for personality traits developed by personality psychologists. Several different psychologists have concluded that personality traits can be organized into five subordinate factors, also called domains known as the Big Five since Goldberg (1971). The five factors (with the acronym OCEAN) are Openness to Experience (also called intellect or culture), Conscientiousness, Extraversion, Agreeableness, and Neuroticism (also called by its obverse, Emotional Stability) and briefly defined as follows:

\[ \text{Openness to Experience} \]
\[ \text{Conscientiousness} \]
\[ \text{Extraversion} \]
\[ \text{Agreeableness} \]
\[ \text{Neuroticism} \]

\[ \text{For a more detailed description, see Costa and McCrae (1992).} \]
• Openness (O) Open individuals are curious and lead a life full of experiences. They are willing to entertain new ideas and unconventional values. They experience both positive and negative emotions more strongly than do closed individuals. People who score low on Openness tend to be conventional in behavior and conservative in outlook.

• Conscientiousness (C) The conscientious individual is goal-oriented, strong-willed, and determined. People high on conscientiousness are scrupulous, punctual, and reliable. Low scorers are less exacting in applying moral principles.

• Extraversion (E) In addition to liking people and preferring company, extraverts are also assertive, active, and talkative. They like excitement and stimulation and are upbeat, energetic, and optimistic. Introverts are reserved and independent.

• Agreeableness (A) The agreeable person is fundamentally altruistic. He or she is sympathetic to others and eager to help them, and believes that others will be equally helpful in return. The disagreeable (also called antagonistic) person is egocentric, skeptical of others’ intentions and competitive rather than cooperative.

• Neuroticism (N) The Neuroticism domain represents the tendency to experience negative affects such as fear, sadness, embarrassment, anger, guilt, and disgust. Men and women high in Neuroticism are also prone to have irrational ideas, to be less able to control their impulses, and to cope more poorly than others with stress. Individuals who score low on Neuroticism are emotionally stable. They are usually even-tempered, and they are able to face stressful situations without becoming upset.

Each of these five factor contains 6 different sub-categories referred to as facets. We use the Danish NEO-PI-R Short Version with 60 items in total that takes most people 10 to 15 minutes to complete. The Danish NEO-PI-R Short Version consists of five 12-item scales measuring each domain. The 12 items for each domain are chosen from the original 48 items as follows: for each facet the two items (out of eight) with the highest correlation with the total factor score are chosen (this is different from the American 60-item version of NEO-PI-R, called NEO-FFI, where the 12 items with the highest correlation with the total factor score is picked, without regard to which facets the single items belong to). In the Danish short version, all facets are therefore represented equally within each domain. The Danish short version is well suited for research where statistical estimations are the main focus. Even though the statistical uncertainty when measuring individual facets will be higher than when using the full NEO-PI-R, the correlations between the facet scores in the short and the long version are high, and the facet scores from the short version can therefore be used when group means are the main focus, c.f. Costa and McCrae (1992). This makes it possible to analyze facets separately.

---

4The questionnaire contains 240 items in total, 48 items to identify each domain. The 48 items can further be divided into six groups of eight that each identifies one facet. Given that the experiment should not last more than 40-50 minutes in total, it would be inappropriate to use a 240-item personality test that takes most people 30-40 minutes to complete.
4.3.9 Intelligence-Structure-Test 2000 R, Matrices

Cognitive ability, or IQ, is measured using the Matrices part of the Intelligence-Structure-Test 2000 R (or I-S-T 2000 R) which is almost identical to Raven’s Progressive Matrices. An advantage of such a test is that it does not depend heavily on verbal skills or other kind of knowledge taught during formal education, cf. Borghans, Duckworth, Heckman, and Weel (2008). In other words, the test is "culture-free". In addition, Millet and Dewitte (2007) find that it also one of the best measures of general intelligence.

The test items can be seen in the screenshots in the appendix. Each of the 20 test items is a matrix presenting a pattern of abstract figures with a missing part. The subject must choose the missing part from a row of five other figures. The subject has 10 minutes to complete the test. Subjects could browse back and forth between the 20 task during the test. The test scores are calculated as the number of correct answers out of 20 possible. The test scores can be converted to conventional IQ scores by multiplying the number of correct answers with a certain fraction. In a German study reported in the manual to the test (Beauducel, Liepmann, Horn, and Brocke 2010), the average test score is 9.6 which corresponds to an IQ of 100. The test can measure IQ’s from about 60 to 150.

In our experiment subjects are not paid for correct answers. The empirical evidence point out that it is unclear whether it is better to pay subjects or not, but the choice made can make a difference to the results. The decision not to pay subjects was partly made to limit the overall cost of payments to subjects in the experiment.

We obtained permission to use the test from Dansk Psykologisk Forlag by agreeing to share the data on the IQ and personality tests and standard socioeconomic variables.

4.3.10 End of first stage

After completing the experiment subjects were informed that they should log in to our webpage again after 1-7 days (depending on when they did the experiment) to see the outcome of the experiment and receive payments (incentivized treatments). To make it harder to hack the payment part, subjects received an additional two-digit control code to use when logging in to the feedback screens.

Subjects was given the opportunity to give us their email if they wanted an email reminder when the feedback part of the experiment opened. Note that this was optional.

5 Payoff information

The payments were transferred to the participants’ bank accounts two weeks after the experiment. The average payment in the incentivized treatments was 276 DKK (∼37 EUR). The lowest payment was 139 DKK (∼18 EUR) and the highest payment was 486 DKK (∼65 EUR).

References


Costa, P., and R. McCrae (1992): *Revised NEO Personality Inventory (NEO PI-R) and NEO Five Factor Inventory (NEO FF-I)*. Psychological Assessment Resources, Odessa, FL.


A Pictures

Figure 3: Letter

Kære [Name]

Danmarks Statistik og Internet Laboratoriet for Eksperimentel Økonomi (ILEE) ved Økonomisk Institut på Københavns Universitet inviterer dig hermed til at deltage i et eksperiment vedrørende økonomiske beslutningsprocesser.

Eksperimenter er et vigtigt redskab inden for økonomisk forskning, idet de er med til at skabe en bedre forståelse for, hvordan mennesker træffer økonomiske beslutninger. I sådanne situationer kan det være vigtigt at forbedre den første økonomiske politik. Et økonomisk eksperiment kan tage mange forskellige former – eksempelvis kan det gå ud på, at deltagerne skal købe og sælge varer på et fiktivt marked eller træffe beslutninger om at investere.

For at opnå et repræsentativt billede har Danmarks Statistik udvalgt et stort antal personer fra hele Danmark, som nu får mulighed for at deltage i eksperimentet. Du er blandt de tilfældigt udtrukne. Dine deltagerstatsdata er naturligvis privat, men vi håber meget, at du vil deltage. Der kræves ingen særlig kendskab til hverken økonomi eller computere for at kunne deltage i eksperimentet, og dine beslutninger i eksperimentet bliver behandlet strengt fortroligt og anonymt.

Ved at deltage i eksperimentet får du mulighed for at tjene penge. Vi kan ikke garantere dig, at du vil tjene et bestemt beløb, idet din indtjening vil afhænge af dine egne samt andre deltageres beslutninger. De nærmere regler er beskrevet på hjemmesiden.

For at sikre deltagerne fuld anonymitet logger alle deltagerer ind med et tilfældigt udvalgt nummer. Vi laver en række forskellige eksperimenter, og alle deltagerer derfor ikke har det samme eksperiment. For at se detaljerne i netop dit eksperiment, herunder opgaven, tidsforbrug mv., bedes du snarest muligt logge ind på vores hjemmeside:

www.econ.ku.dk/ilee med dit login nummer: 28.826-6

Hvis du har problemer med at logge ind eller har yderligere spørgsmål, er du velkommen til at kontakte Økonomisk Institut på e-mail ilee@econ.ku.dk eller telefon 35 32 44 09.

Med venlig hilsen og på forhånd tak for din hjælp.

Isak Ilksen
Kontorchef, Danmarks Statistik

Jean-Robert Tyran
Professor, Økonomisk Institut

Danmarks Statistik
Sejnegade 11
2100 København Ø
Tlf. 39 17 39 17
Fax 39 17 39 99
CVR 17-15-04-13
dst@dst.dk
www.dst.dk
Figure 5: Gift
B  Instructions and Screenshots

Complete translated screenshots are available in the following documents:

Main Experiment:
- Paid Give Translated Screenshots - Main Experiment.pdf
- Hypothetical Give Translated Screenshots - Main Experiment.pdf
- Paid Take Translated Screenshots - Main Experiment.pdf

Feedback Part:
- Paid Give Translated Screenshots - Feedback Part.pdf
- Hypothetical Give Translated Screenshots - Feedback Part.pdf
- Paid Take Translated Screenshots - Feedback Part.pdf

C  Invitation Letters

The invitation letters in Danish and with English translations are contained on the following pages.