Replication of

Lab Experiments for the Study of Social-Ecological Systems

by Janssen, M. A. / Holahan, R. / Lee, A. / Ostrom, E. (2010)
in: Science, 328, pp. 613–617

Replication Authors:
Teck-Hua Ho and Hang Wu

In a common-pool resource game with spatial and temporal resource dynamics, Janssen et al. (2010) tested the effects of costly punishments and communication. Subjects were randomly assigned to six different treatments. Each treatment consisted of three consecutive 4-minute periods of costly punishment (P), communication (C), or a combination of both (CP), and three consecutive 4-min periods when neither communication nor punishment (NCP) was allowed. All treatments thus consisted of six decision periods, each lasting 4 minutes. Half of the treatments started with NCP and the other half finished with NCP.

One of the key findings is that in the first three periods of the experiment, if the experiment started with the C condition, then average earnings increased as compared to if the experiment started with the NCP condition. This finding will be the focus of our replication; we only include the treatments that started with the NCP and C conditions.

Hypothesis to replicate and bet on:

Communication increases average earnings in a common-pool resource game with spatial and temporal resource dynamics. A comparison of net earnings between the *NCP* condition and the *C* condition in periods 1 to 3 showed p-value < 0.001 with the Mann-Whitney test ($z = 5.761$ and $p = 8.362e^{-9}$).

Power Analysis and Criteria for Replication: First Data Collection

The original sample size included 63 group-and-period observations (21 groups of 5 individuals observed in 3 periods) and 105 individuals; the standardized effect size measured as the correlation coefficient (r) was 0.631. In order to have 90% power to detect 75% of the original effect size, a sample size of 42 group-and-period observations (70 individuals) is required. The criterion for replication is an effect in the same direction as the original study with a p-value < 0.05 (Mann-Whitney test).

Power Analysis and Criteria for Replication: Second Data Collection

If the original result is not replicated in the first data collection, a second data collection of 63 additional group-and-period observations (105 individuals) will be carried out so that the total sample size is 105 group and period observations (175 individuals).

If a second data collection is carried out, we will also test whether the original result replicates in the pooled sample of the first and second data collections.

In order to have 90% power to detect

50% of the original effect size, a sample size of 104 group-and-period observations is required. But as this number needs to be evenly dividable by three (due to the three periods) we will collect 105 group-and-period observations in total if a second data collection is needed (and thus 63 rather than 62 group-and-period observation in the second data collection). The criterion for replication is an effect in the same direction as the original and a p-value < 0.05 (Mann-Whitney test) in the pooled data.

Sample

The sample size in the first data collection consists of 42 group-and-period observations (70 individuals) from the National University of Singapore (NUS).

If the original result is not replicated in the first data collection (Mann-Whitney test with a p-value < 0.05 in the original direction), a second data collection of 63 group-and-period observations (105 additional individuals) from NUS will be carried out so that the total number of group-and-period observations is 105 (175 individuals).

Materials

We use the same computer program as used in the original article, provided by the original authors (open-source software also available at http://commons.asu.edu).

Procedure

We follow the procedure of the original article. Subjects will be recruited through recruiting advertisements posted in the NUS campus, as well as e-mail invitations sent to an existing voluntary database of undergraduate students maintained by the Centre for Behavioural Economics (CBE) at NUS. The following summary of the experimental pro-

cedure is therefore based on section 1.2 (pp. 4–6) of the Supplementary Information.

Two days before a session was conducted, a recruitment email was sent to members of the database informing them of an opportunity to participate in an experiment. Potential participants were informed that the experiment should last about 90 minutes, that they would be guaranteed a $10.00 show-up payment, and that they should expect to earn in the neighbourhood of $15.00–$30.00 for their participation of about 60 minutes, but that the exact amount of earnings would depend on the results of the experiment. If a participant chose to sign up, (s)he clicked on a link in the email that took him/her to a secure website that listed all available experimental sessions. For each session we recruited 18 participants, three of whom became alternates.

Experiments included either 10 or 15 participants, depending on turnout. When participants arrived at the lab they were greeted by an experimenter and asked to read and sign an informed consent form while they waited for more participants to arrive. Our experiment required groups of five people and, in order to maintain some degree of anonymity, we only ran sessions if enough participants to fill either 2 or 3 groups showed up. Once a sufficient number of participants arrived, they were escorted into the computer room. Alternates were randomly selected through a lottery, paid $10.00, and asked to leave.

After everyone sat down at a computer terminal, the experimenter introduced himself and explained that they would be participating in a real-time group decision-making experiment with other participants in the room. At that point participants were asked to remain silent for the duration of the experiment, to turn off cell phones, and to put all other work away. The computer room at the lab consists of individual computer terminals, separated by dividers to prevent participants

from seeing one another's screens. When the experimental software was loaded, it randomly assigned each computer terminal to a group and gave each participant an identification number that remained constant throughout the experiment. In this way members of a group could gauge an individual's over-time behavior without actually knowing which participant in the room had been assigned which identification number. Once everyone was settled in and identification numbers had been assigned, the general instructions were read.

Our experiments use a real-time renewable resource environment. Participants see themselves on screen as a yellow avatar with the ability to collect green tokens by moving over the token and pressing the space bar on their keyboard (Fig. S3 in Supplementary Materials). The experimental environment requires some practice and we ask participants to answer two questions about the experimental environment after the general instructions are read to ensure they understand some key aspects of the regeneration of the resource. After that, participants participate in a four-minute practice period, in which they are individually placed in a 13×13 grid of cells to practice moving their avatars and collecting tokens. During the practice period, participants can reset the distribution of tokens by pressing the R key, so they can continue to practice even if they collect all the tokens on the screen before the four minutes expire.

At the completion of the practice period, participants were informed that the experimental environment would become five times larger and would be shared among the five members of the group. In addition, participants were told that they could no longer reset the distribution of tokens. Then the instructions for the first period were read. Since there are six treatments, the instructions for the first period varied slightly across treatments. After the first period was completed,

participants were simply informed that the instructions for the next period were identical to the first period and the same was said before the start of period 3. Before the fourth period, however, new instructions were read to correspond with the experimental treatment. Participants were informed during the reading of the general instructions that the experiment would be six periods long, but they were not informed at any point that the experiment consisted of two stages of three periods each. Instead, they were only given instructions by period.

After the completion of the sixth period, participants were asked to fill out a brief survey that asked for basic demographic information and satisfaction with the experiment. While participants filled out the survey, one experimenter stayed in the computer lab, while another experimenter prepared participant payments in the reception area. Participants were then asked to leave one-by-one and sign for their cash payments. In this way, participants were not able to learn about the earnings of other participants.

Analysis

The analysis will be performed exactly as in the original article. In the original article, a Mann-Whitney test was used to test the equality of distributions of earnings between the treatment starting with the *NCP* condition and the one starting with the *C* condition in the first three periods (Table S4B in the Supplementary Information); $U = 717(48, 15)$ and $z = 5.761$. The same test will be used in the replication.

The results will first be estimated based on the first data collection. If the original result is replicated in the first data collection (a two-sided p-value < 0.05 in the same direction as the original study), the second data collection will not be carried out.

If the original result is not replicated in the first data collection a second data collection will be carried out. The above statistical test will then be estimated for the pooled sample of the first and second data collections to test if the original result replicated (a two-sided *p*-value < 0.05 in the same direction as the original study).

The experiment will be in English, as was the original study.

Differences from Original Study

The replication procedure is the same as that used in the original study, with some unavoidable deviations. The replication will be carried out in NUS between September 2016 and September 2017, whereas the original experiment was carried out at Indiana University-Bloomington in 2008.

The original paper contains six treatments. Three treatments started with the *NCP* condition (*NCP-P*, *NCP-CP*, and *NCP-C*), one treatments started with the *C* condition (*C-NCP*), and two treatments started with other conditions. We only replicate the *NCP-C* and the *C-NCP* treatments for all six periods even though only data for the first three periods are used in the analysis.

In the original experiments participants were paid $10.00 show-up fee plus $15.00 to $30.00 performance-based payment. In the replication, the participants will be paid in Singapore dollars using the same payment scheme, i.e., s$10.00 show-up fee plus s$15.00 to s$30.00 performance-based payment.

Replication Results for the First Data Collection (90% power to detect 75% of the original effect size)

[To be added when replication experiments have been completed.]

Replication Results for the First and Second Data Collection Pooled (90% power to detect 50% of the original effect size)

[To be added when replication experiments have been completed.]

Unplanned Protocol Deviations

[To be added when replication experiments have been completed.]

Discussion

[To be added when replication experiments have been completed.]

References

Janssen, M. A. / Holahan, R. / Lee, A. / Ostrom, E. (2010): *Lab experiments for the study of social-ecological systems*, Science, 328, pp. 613–617.