*Replication of*

# Cooperating with the Future

by Hauser, O. P. / Rand, D. G. / Peysakhovich, A. / Nowak, M. A. (2014)
in: Nature, 511, pp. 220–222

***Replication Authors:***
*Adam Altmejd, Anna Dreber, Emma Heikensten, Siri Isaksson, and Magnus Johannesson*

---

*Hauser et al (2014) let subjects of Amazon Mechanical Turk (AMT) play an Intergenerational Goods Game (IGG) with three different institutional treatments (unregulated, voting and partial voting). The IGG builds on the public goods game, common pool resource games and threshold games. Subjects are assigned to groups of five, referred to as generations and the first generation is endowed with 100 units. Each group member makes a choice on how much to extract, between 0 and 20 units. If less than 50 units are extracted by the group collectively, the pool of units will renew to 100 units for the next generation and thus be sustained. If more than 50 units are extracted, the pool will be exhausted and subsequent generations will not receive anything. The pool of resources is renewed and a new generation occurs with the probability of 0.8 and the game ends with a probability of 0.2. In the unregulated treatment, the subjects are asked to individually choose how many units to extract from the common pool of units. In the (full) voting treatment, the group members are able to vote on how much they think that the group should extract and the median vote is automatically extracted by each group member. Similarly, in the partial voting treatment only three out of five subjects vote and are bound to follow the median vote among themselves. The other two are not told about the results from the voting and choose individually how much to extract. In addition to these treatment variations, the authors also test the effects of different threshold levels and different probabilities of a new generation occurring in the voting treatment. In total there were seven experimental conditions and the subjects were randomly assigned to each one.*

**Hypothesis to replicate and bet on:**

Choosing an extraction level for all group members using median voting leads to a higher degree of sustainability of a common pool than allowing each individual to choose their own extraction amount. That is, a comparison of the average probability that the common pool was sustained by the first generation between the voting treatment and the unregulated treatment (in both treatments there is an 80% probability that a new generation occurs and an extraction threshold of 50%). To evaluate this hypothesis, a linear probability model with a treatment dummy variable is used; see the 1st generation regression equation in Table S1; $p = 1.427e^{-10}$ (reported as $p < 0.001$) in a $t$-test ($t(38) = 8.696$) of the treatment dummy variable coefficient.

## Power Analysis and Criteria for Replication: First Data Collection

The original sample size was 40 pools where one pool is a game with expected length of 5 generations and 5 players in each generation, where only the first generation is included in the test. The standardized effect size measured as the correlation coefficient ($r$) was 0.816. To have 90% power to detect 75% of the original effect size a sample size of 22 pools is required. The criteria for replication is an effect in the same direction as the original study and a $p$-value $< 0.05$ (in a two-sided 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 38 pools will be carried out so that the total sample size is 60 pools. If a second data collection is carried out, it will be tested if the original result replicates in the pooled sample of the first and second data collection.

To have 90% power to detect 50% of the original effect size a sample size of 59 pools is required; i.e. an additional sample size of 37 pools in the second data collection to arrive at a sample size of 59 pools in total for the first and second data collection pooled. If a second data collection is carried out, we collect 38 pools rather than 37 in the second data collection to have the same number of pools in both treatments. The criteria for replication is an effect in the same direction as the original and a $p$-value $< 0.05$ (in a two-sided test) in the pooled data.

## Sample

The sample in the first data collection consists of 22 pools recruited from AMT. If the original result is not replicated in the first data collection (two-sided $p$-value $< 0.05$ in the original direction) a second data collection of 38 additional pools from AMT will be carried out so that the total sample size is 60 pools.

## Materials

We use the same computer program as used in the original article. The instructions will be the same with the exception that a consent form is added to the replication (a consent form was not included in the original study).

## Procedure

The original paper contains seven experimental conditions/treatments: different institutions (unregulated, voting, or partial voting), different values of the probability of a new generation occurring (the game continuing) and the threshold for exhausting the common pool. For the replication we only replicate two of these treatments: the unregulated treatment and the voting treatment with a probability of a new generation occurring of 80% and a threshold of exhausting the common pool of 50%.

In the unregulated treatment, each of the five group members independently selected an extraction amount between 0 and 20 units. Under the voting institution, each of the five group members proposed an extraction amount between 0 and 20 units. The median proposal amount was then extracted for each group member.

We follow the procedure of the original article (with the exception of the higher compensation for completing the HIT, and that we only implement the two experimental treatments mentioned above). Subjects will be recruited using AMT, with a new requester account, and will receive a $1.00 show-up fee (rather than $0.50 as used in the original study) for participating, with a chance to earn

up to an additional $1.00 based on the outcome of the experiment. If necessary to recruit a sufficient number of participants the $1.00 show-up fee will be increased.

The following summary of the experimental procedure is based on section 2 (p. 4–6) of the Supplementary Information. The participants will take part of the experiment through an online survey provided by Qualtrics. The subjects will be randomly allocated to a treatment by an internal randomization mechanism of Qualtrics.

As the experiment starts subjects read instructions and are then asked to complete a comprehension quiz. If they pass the test they are asked to continue. If they don't pass the quiz they receive only the baseline payment of $1.00, and are excluded from participating. Otherwise they continued to the main decision making stage. After the main decision making stage the subjects get to read the instructions for the partial voting treatment and take a second comprehension quiz. Only participants that passes both quizzes will be included in the data.

Generations are recruited sequentially and each generation is informed of the binary outcome of the previous generation (but not the extraction amounts). Once a game is discontinued (i.e., based on the 80% probability of the game continuing to the next generation), no more groups are recruited. Participants in Generation 1 are informed that they are the first generation but participants in subsequent generations are not informed of their specific generation number (other than showing that they were not the first generation).

Within each generation of a game, a group of five participants choose how many units to extract from the pool of 100 units (in the unregulated treatment each individual chooses her extraction level and in the voting treatment the median proposal is extracted for each group member). If the fraction extracted within a given group does not exceed the extraction threshold, that group's pool is 'sustained' and the next generation receives a pool refilled to 100 units and has a chance to make their own set of extraction decisions (provided that the game was randomly drawn to continue for another generation based on the 80% continuation probability). If the fraction extracted exceeded the threshold, the pool is exhausted. All future generations are then informed that a previous generation extracted more than the threshold and as a result they don't have the opportunity to play the IGG or receive any bonus payment.

Once the decisions had been made by all members of a generation (the group of five), payoffs were calculated and the subjects were paid a $1.00 show-up fee and their payoffs in the IGG through AMT (each subject's extraction level was converted into dollars at the exchange rate of $1.00 = 20 units of the common pool).

The randomization and matching processes that were used in the original experiment will also be used in the replication. The number of generations in each pool will be predetermined given the 80% continuation probability. Then, for each generation the program will randomly sort the subjects into one of the two treatments. After all subjects in one generation have completed the tasks and control questions, we will match subjects within the same treatment that answered the control questions correctly into groups of five. If the number of subjects is not divisible by five (the size of the group) some will be chosen randomly to not be included in the analysis, although they will still be paid in accordance with their response. The experiment will be in English as in the original study.

## Analysis

The analysis will be performed exactly as in the original article. The following summary of

the analysis is therefore based on section 2.5 (p. 7) and section 4 (pp. 8–9) of the Supplementary Information.

Results are analyzed in a linear probability model with the probability that the pool is sustained by the 1st generation as the dependent variable (first generation groups that did not extract more than 50 units will be coded as 1 and first generation groups that extracted more than 50 units will be coded as 0) and a "voting treatment" dummy variable as the independent variable (the regression equation for the 1st Generation in Table S1 in the Supplementary Information). Robust standard errors will be estimated.

In the original article the "voting treatment" dummy variable had a coefficient of 0.800 ($SE = 0.092$) and a $p$-value of $1.427e^{-10}$ (reported as $p < 0.001$) based on a $t$-test of the regression coefficient ($t = 8.696$). The same test will be used in the replication.

The result for the 1st generation 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 for the 1st generation will then be estimated for the pooled sample of the first and second data collection to test if the original result replicated (a two-sided $p$-value $< 0.05$ in the same direction as the original study).

The result for the 1st generation above will be the main replication result. However, for completeness, we will also report the result for all generations in the original article. This result is estimated with a linear probability model with the probability that the pool is sustained in each generation as the dependent variable (groups that did not extract more than 50 units will be coded as 0, while any groups extracting more than 50 units will be coded as 1; if the pool is not sustained the variable is coded as 0 for all groups in future generations) and a "voting treatment" dummy variable as the independent variable (the regression equation for All Generations in Table S1 in the Supplementary Information). Robust standard errors will be estimated (according to the original authors standard errors were not estimated with clustering on the pool; but in the replication we will report the result both with and without clustering on the pool).

## Differences from Original Study

The replication procedure is the same as that of the original study, with some deviations. The replication will be performed at AMT between September 2016 and September 2017, whereas the data in the original study was carried out at AMT in 2014. A consent form is also added to the replication. The participation payment (show-up fee) for completing the HIT was \$0.50 in the original study and will be \$1.00 in the replication, to be able to recruit a sufficient number of participants to the study. If necessary to recruit a sufficient number of participants the \$1.00 show-up fee will be increased. Participants who participated in the original study will be excluded from participating in the replication as suggested by the original authors (based on the list of Turker Id's of previous participants provided by the original authors).

The original paper contains seven experimental conditions/treatments. For the replication we only replicate two of these treatments: the unregulated treatment and the voting treatment with a probability of a new generation occurring of 80% and a threshold of exhausting the common pool of 50%.

## 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

Hauser, O. P. / Rand, D. G. / Peysakhovich, A. / Nowak, M. A. (2014): *Cooperating with the future*, Nature, 511, pp. 220–223.