

Replication of

The Potential of Social Identity for Equilibrium Selection

by Chen, R./Chen, Y. (2011)

in: The American Economic Review, 101(6), 2562–2589.

Replication Authors:

Teck Hua Ho, Hang Wu

Chen and Chen study effort in a minimum effort game varying whether subjects have a salient ingroup or outgroup identity. They find that effort is higher for those with a salient ingroup identity.

Hypothesis to bet on:

Effort in a minimum effort game is higher for subjects with a salient ingroup identity than for subjects with a salient outgroup identity (a comparison of mean effort between the Enhanced Ingroup treatment and the Enhanced Outgroup treatment).

Power Analysis

The original p -value is 0.023 (a comparison of the sum of the regression coefficients “Ingroup” and “Ingroup * Enhanced” versus “Outgroup” and “Outgroup * Enhanced” in the first column in Table 2 and on page 2577 (random effects regression, with clustering at the session level); this tests the difference between the Enhanced Ingroup treatment and the Enhanced Outgroup treatment; in the replication only these two treatment needs to be included and the difference tested with a single treatment dummy): “A test that the sum of the coefficients on the ingroup dummy and ingroup-enhanced interaction term is equal to 0 yields $p < 0.0001$ for column 1 and $p = 0.0005$ for column 2, while a test that the previous sum is equal to the corresponding outgroup sum yields $p = 0.023$ for column 1”

As this p -value is based on a regression model including more than the two treatments in the replication, we re-estimated the regression model including only the two treatments of the replication. This yielded a p -value of 0.033;

based on a χ^2 test of the dummy variable for the Enhanced Ingroup treatment (the original article also tested the hypothesis with a χ^2 test).

The original sample size is 72 participants (36 in the Enhanced Ingroup treatment and 36 in the Enhanced Outgroup treatment). To achieve 90% power the required sample size is 166 participants.

Sample

The sample for replication consists of 168 students (84 participants in each treatment) at the National University of Singapore. There are no exclusion criteria.

Materials

We use the material of the original experiment (programmed in z-Tree) along with the original instructions, both available on the journal’s webpage.

Procedure

We follow the procedure of the original article, with only slight but unavoidable deviations as outlined below. The following summary of the experimental procedure is therefore based on the section “III.B. Experimental Procedure” (pp. 2570–2573) in the original study.

12 subjects in a session are randomly divided into two groups. Each subject randomly chooses an envelope which contains a red or a green index card with a subject ID number on it. Each subject is assigned to the red or the green group based on this index card; each group has 6 members.

Subjects are then given five minutes to review five pairs of paintings, each of which contains one painting by Paul Klee and one painting by Wassily Kandinsky. Subjects are then shown two final paintings and are told that each of them was painted by either Klee or Kandinsky, and that they both could have been painted by the same artist. The subjects are asked to determine, within ten minutes, which artist painted each of these final two paintings. Each subject is allowed to use an online communication program to discuss the problem with other members of her own group. A subject is not required to give answers that conform to any decision reached by her group, and she is not required to contribute to the discussion. For each correct answer, a subject earns 350 tokens (the equivalent of \$1), though she is not told what the correct responses are until the end of the experiment.

Each subject then plays a minimum-effort game for 50 rounds. For each round, each subject is randomly rematched with one other subject in the same session. In the Enhanced Ingroup treatment subjects are matched only with members of their own group, and in the Enhanced Outgroup treatment subjects are matched only with members of the other group.

At the end of the experiment subjects fill out a post experimental survey which contains questions about demographics, past giving behaviour, strategies used during the experiment,

group affiliation, and prior knowledge about the artists and paintings.

In the original study, the experiment was carried out in groups of 12 subjects per session. We will include one group of 12 subjects from each of the two treatments in each session (i.e. 24 subjects per session and 168 subjects in total). Subjects will be randomly allocated to the two treatments within each session.

After all rounds have been played, subjects will be privately paid in cash using the same show-up fee (\$5) and incentives as in the original study (average earnings were \$11.69 per subject in the original study).

Analysis

The analysis will be performed exactly as in the original article using the random effects regression specification without demographic variables included.

Differences from Original Study

The replication procedure is identical to that of the original study, with some unavoidable deviations. This replication will be performed at the National University of Singapore in Singapore, in 2015, on students from the National University of Singapore, while the original data was gathered at the University of Michigan in Ann Arbor, USA, in 2007-2008, on students from the University of Michigan (plus one student from Eastern Michigan University and one non-student). The experiment will be in English as in the original study.

The original study includes a control treatment and also three “Near-minimal” treatments: for the replication the focus is only on the difference between the Enhanced Ingroup treatment and the Enhanced Outgroup treatment so only these two treatments will be included.

Replication Results

In the replication experiment, though the effect was in the expected direction, the coeffi-

cient of 5.204 is not statistically significant with a p -value of 0.571 (see Table 1). In comparison to the effect size of 23.852 in the original study, the relative effect size of the replication experiment equals 21.82% (5.204/23.852).

Figure 1 presents the mean and median effort dynamics observed in the two treatments of the replication experiment. One observation is that the outgroup sessions started with higher efforts than the ingroup sessions, indicating a lack of ingroup favoritism effect at the beginning. However, group identity seems to have an influence on the effort levels over time. The mean effort level in the outgroup sessions move down from 147.274 in the first period to 134.253 at the end, which is quite a substantial decrease as compared to the minor decrease (from 141.857 to 140.894) in the ingroup treatment. This difference is even more pronounced for the time series of median effort levels.

Unplanned Protocol Deviations

In the above protocol, we planned to include 24 subjects for each session and randomly allocate subjects into two groups with 12 subjects for each group upon arrival. In the replication experiments we followed the procedures used in the original study and included only 12 subjects for each session (in total 14 sessions with 7 sessions for each treatment). Besides, there was a slight difference between the materials used in our replication and the original experiment: for one painting (1A *Gebirgsbildung*, 1924, by

Klee), we used a black and white print of the painting while the one used in the original experiment was in color. Apart from that the replication experiments were conducted exactly in the way as described above without any deviations from the protocol.

Discussion

Given the criteria and procedure outlined above, the hypothesis of interest has not been replicated at a significance level of $\alpha < 5\%$. The relative effect size equals 21.82% and the p -value of the hypothesis test is 0.571.

The key difference between the replication and the original experiment is that the replication used Singapore rather American university student sample. One potential reason for the lack of an ingroup favoritism effect in the replication study could be that Singapore subjects were less active in the communication stage than the American subjects, and, hence, the group identity was less salient as compared to the original study. In the replication experiment the average number of chat lines of a group in the communication stage was 44, which is significantly less than the original study for which the number was 64. This difference could have played a role. But it does not seem likely to be a critical one, as more communication (in terms of number of chat lines) only has a positive but not significant effect (coefficient = 0.188, $p = 0.180$) on the effort level.

Table 1: Comparison of random effects regressions of effort choice on group identity in both treatments

| | <i>Original Study</i> | <i>Replication Study</i> |
|-----------------------|-----------------------------|--------------------------|
| 1 if Enhanced Ingroup | 23.852** (11.200) | 5.204 (9.197) |
| Constant | 139.482*** (11.072) | 139.156*** (7.459) |
| Observations | 3600 | 8400 |
| R^2 | 0.292 | 0.012 |

Note: Standard errors adjusted for clustering at the session level are shown in parentheses.

- *** Significant at the 1 percent level
- ** Significant at the 5 percent level
- * Significant at the 10 percent level

Figure 1: Mean (left) and median (right) effort dynamics in the replication experiment

