

Replication of Cooney et al 2016
“When fairness matters less than we expect”
PNAS 2016, 113 (40), 11168-11171.
<https://www.pnas.org/content/113/40/11168>

The original paper includes several studies. We randomly chose study 5. In this within-subject experiment, participants are recruited online and given instructions to a modified dictator game with three participants: an allocator and two receivers. All participants are told that they are assigned to be receivers (note: the experiment is deceptive, as the receivers are made to believe that they are in an actual dictator game). They are told that an allocator will decide to give a bonus to one of two receivers, and the allocation method is randomly assigned to the allocator: in the fair procedure, the allocation is decided using a fair coin flip while in the unfair procedure the allocation procedure is determined by the allocator. Before the allocation is made, participants predict how they will feel about being or not being picked to receive the bonus depending on the procedure on a 7 point Likert-scale. Participants expect that fairness will have an impact on how they would feel, in that they would feel less bad not receiving the bonus under the fair procedure compared to not receiving the bonus under the unfair procedure.

Hypothesis to replicate and bet on: Participants predict that they would feel less bad not receiving the bonus under the fair procedure compared to not receiving the bonus under the unfair procedure. Participants are asked to predict how they would feel under four different scenarios: 2 (Outcome: receive bonus or no bonus) x 2 (Procedure: fair or unfair). To evaluate this hypothesis, the authors perform an F -test on the Outcome x Procedure interaction ($F(1,119) = 54.23, p < 0.001, \eta^2_p = 0.313$); p. 4 in SI.

Criteria for replication: The criteria for replication are an effect in the same direction as the original study and a p -value < 0.05 in a two-sided F -test.

Power analysis: The original study had 120 participants. The standardized effect size (Cohen's d) was $d = 0.672$. To have 90% power to detect 67% of the original effect size, a sample size of $n = 53$ is required. Since we require the replication sample size to be at least as large as in the original study, the replication experiment will use a sample size of $n = 120$ and the power will thus be $>90\%$ to detect 67% of the original effect size.

Sample: The original paper mentions no restrictions on who could participate, and no observations were excluded. We will make sure that participants can only participate once from the same account in this specific study, and we will only recruit participants with a HIT approval rate of 95% or above. We will also check all IP addresses via <https://www.ipqualityscore.com/>; and we will remove any participants where one or more of the following is true: fraud score ≥ 85 ; TOR = True; VPN = True; Bot = True; abuse velocity = high. The replication sample size is the sample size after any exclusions of participants.

Materials: We will use the same material as in the original study, kindly provided by the original authors. In particular, the experiment will be conducted using the original *Qualtrics* survey, including the same instructions and images as in the original study.

Procedure: We will closely follow the procedure of the original experiment. The following summary of the experimental procedure is therefore largely based on the description of the experiment in the article (p. 11170) and the Supporting Information (pp. 3–4).

Participants will first be shown a Captcha, and will thereafter provide informed consent. After this we will include an attention check that participants will need to pass to continue to the study. This attention check is in addition to any other potential attention check(s) used in the original study. Participants will then be told that the survey involves interactions with two other participants and that they should wait for two other participants to join their group. A ticking clock will appear on the screen, and after 15 seconds the page automatically advances. However, there are no other participants but to ensure that participants believe that they are interacting with other participants, bogus participant identification numbers will be presented at the beginning of the study and at various other points during the study. Participants will, at various points in the study, also be asked to wait so that the other participants can “catch up.” Participants will be told that one of them will be randomly assigned to play the role of decider and that the others will be assigned to play the roles of receivers.

After this, participants will be told about the modified dictator game, including that the decider will be assigned to use one of two procedures to allocate a bonus to one and only one of the two receivers. Participants are told that deciders will be required to allocate the bonus either (i) by “picking,” which involves selecting one of the receivers to receive the bonus and the other to receive nothing, or (ii) by “flipping,” which involves flipping a “digital coin” to determine which of the receivers will receive the bonus and which will receive nothing. The bonus is set at \$0.25.

Participants will then be told that they have been randomly assigned to play the role of a receiver. Before the allocation is made, receivers will be asked to answer the questions “How would you feel if the decider is asked to pick to determine who will receive \$0.25 and who will receive nothing?” and “How would you feel if the decider is asked to flip a coin to determine who will receive \$0.25 and who will receive nothing?” Receivers will answer two versions of each question: one that was followed by the phrase “if you receive \$0.25” and one that was followed by the phrase “if you receive nothing.” Receivers answer these four questions by clicking on a series of 7-point Likert scales whose endpoints are labeled *not very good* (1) and *very good* (7). After receivers made their predictions, they will be informed that the study is over and they will be awarded \$0.25 as “a thank you for your time.” Receivers will then complete a series of demographic questions.

Analysis: The analysis will be performed as in the original paper. The analysis code was kindly provided by the original authors. Specifically, we will perform a 2 (Outcome: bonus or no bonus) \times 2 (Procedure: fair or unfair) ANOVA. The replication focuses on the *F*-test on the Outcome \times Procedure interaction.

Subject payments: We are standardizing payments across all replications so that studies have a certain show-up fee depending on the expected length of the study, with an hourly wage from the show-up fee of \$8 and a minimum payment of \$1 (for studies with incentive payment we use the same incentive payment as in the original study; and this payment is paid in addition to the show-up fee). If we have problems recruiting, we will increase the show-up fee.