Replication Protocol Draft
Study 7 from Rand, Greene, & Nowak (2012)

Article Citation: Rand, D. G., Greene, J. D., & Nowak, M. A. (2012) Spontaneous giving and calculated greed. *Nature, 489*, 427–430.

Rationale: Rand, Greene, and Nowak (2012) argued that our social intuitions are shaped by our daily experiences, and that those intuitions can determine whether our default response is selfish or cooperative. According to this social heuristic hypothesis, people who have cooperation rewarded in daily life will tend to develop cooperative intuitions as a default response, and those who have non-cooperation rewarded will tend toward selfish intuitive responses. Although intuitions vary across people, deliberation is theorized to always favor self-interested behavior. For example, in one-shot, anonymous economic games, where selfish actions maximize one's payoff, deliberation will favor non-cooperation, overriding any potential intuitive bias toward cooperation. Under these conditions, the social heuristic hypothesis predicts more cooperation for judgments made intuitively (because some participants will default to cooperative responses) than when judgments are made with more deliberation (because deliberation will favor selfishness for all participants).

Some of the central empirical support for this prediction of the social heuristic hypothesis comes from two experiments in Rand et al's (2012) paper (Study 6, an internet experiment with Amazon Mechanical Turk participants, and Study 7, an experiment in an economics laboratory with college students) that revealed a causal relationship between time pressure and cooperative decisions in a one-shot public goods game. In each experiment, half of the participants were required to decide how much to contribute to the group within 10 seconds (time pressure condition/intuitive decision making), whereas the other half of the participants had to think and reflect for at least 10 seconds before deciding on their contribution (reflection/forced delay condition). In both experiments, the mean amount contributed to the public good was

greater in the intuition/time-pressure condition than in the reflection/forced-delay condition, henceforth termed the intuitive-cooperation effect.

Since its publication, Rand et al's (2012) paper has been highly influential, with more than 200 citations in Google scholar. Some studies have replicated the original finding (e.g., Cone and Rand, 2014; Rand and Kraft-Todd, 2014); Rand, Newman, and Wurzbacher, 2015); Rand et al., 2014), and a recent meta-analysis including both published and unpublished direct replication studies found support for the time pressure effect (Rand et al., 2014). Conceptual replications have manipulated the reliance on intuition versus deliberation in public goods games using a writing induction (Lotz, 2014; Rand et al., 2012), ego depletion (De Dreu, Dussel, & Ten Velden, 2015), and cognitive load (Døssing, Piovesan, & Wengstrom, In prep). Other studies have observed positive effects of cognitive load on prosociality in zero-sum monetary transfers (e.g. the Dictator Game; Cornelissen, Dewitte, & Warlop, 2011; Roch, Lane, Samuelson, Allison, & Dent, 2000; Schulz, Fischbacher, Thöni, & Utikal, 2014).

However, the size of the effect of time pressure on cooperation was observed to decline over a two year period and eventually disappear in studies with Mechanical Turk participants (e.g., Rand et al., 2014; Verkoeijen and Bouwmeester, 2014). Another set of three attempts to replicate Study 6 from Rand et al (2012) did not observe a statistically significant intuitive-cooperation effect in a sample of Austrian students or in two general population samples in the United States (a Decision Research Sample and a Qualtrics Panels sample; Tinghög et al., 2013).

Rand et al (2014) suggest that part of this variability results from experience with these sorts of decisions: Participants who have completed economic games of this sort have had a chance to learn that their intuitions can lead them astray in 1-shot games, and thus are less likely to show the intuitive cooperation effect. The declining size of the effect on MTurk could be due to MTurk participants becoming increasingly experienced with economic games as MTurk has become popular for psychological studies (see Rand & Kraft-Todd, 2014 and Dreber, Fudenberg, Levine, & Rand, 2014, for evidence that more experienced MTurk workers put les faith in their intuitions). Consistent with this explanation, Rand et al (2012, Study 9) showed that the intuitive-cooperation effect emerged for naïve participants, but not for experienced participants (using a writing

induction to promote reliance on intuition versus deliberation). This moderating effect of experience was conceptually replicated using time constraints to manipulate cognitive processing (Rand et al, 2014; Rand & Kraft-Todd, 2014).

The goal of this Registered Replication Report is to assess the size and variability of the intuitive-cooperation effect with naïve participants. More specifically, it will measure the size and variability of the intuitive-cooperation effect reported by Rand et al (2012), among naïve college students in a laboratory setting by replicating the between-subjects comparison (time pressure vs. forced delay) from Study 7.

Materials: The first author of the original study provided the original study materials, including the program used to run the original procedure, the instructions, and the post-experimental questions on task understanding and experience.

Design: The proposed study will have to two between-subjects conditions: Time Pressure and Forced Delay. In the Time Pressure condition, subjects will be asked to decide on a contribution to a one-shot public goods game within 10 seconds. In the Forced Delay condition, they will be asked to wait at least 10 seconds before they make their contribution. Each participant's contribution as well as their actual decision time will be recorded. After the primary task, we will include questions designed to measure the following: (a) comprehension of the task, (b) justification for their contribution, (c) individualism/collectivism (Singelis, Triandis, Bhawuk, & Gelfand, 1995), (d) experience with tasks of this sort, (e) experience with research participation more generally, (f) self-reported perceptions of trust in others (a factor thought to moderate the intuitive-cooperation effect), (g) awareness of the research hypothesis (PARH: cf. Rubin, Paolini, & Crisp, 2010), (h) sex, age, and country, and (i) how many of the participants in the room they know.

Critical result: The critical effect size is based on the percentage of the money that participants contributed to the public good. Specifically, for each participant, the amount they contributed is divided by the total possible contribution, yielding the proportion contributed (which is then multiplied by 100 to convert to a percentage). Each lab will

report the average percentage in the time pressure condition, the average percentage in the time delay condition, and the difference between the two measures (i.e., the "risk difference"). In the original study, the mean percentage of funds donated in the one-shot public goods game was higher in the time-pressure condition ($M = 58\%$) than in the forced-delay condition ($M = 42\%$), a risk difference of 16%. A re-analysis of this original study, combining data from the lab and MTurk and including subjects who disobeyed the time constraint, suggested a smaller effect (~6%; Rand et al, 2013). And, a meta-analysis of 15 studies (Rand et al, 2014) observed a risk difference of 3.6% when including subjects who disobeyed the time constraint. These analyses suggest that the observed effect might be smaller than originally reported.

Protocol Requirements

Sample size: n ≥ 150 total, with n ≥ 75 in each of the two conditions. Note that with this sample size, individual studies would be underpowered to reject the null hypothesis for the original effect size. We encourage laboratories to test as large a sample as possible. We set this target sample size with the goal of encouraging participation by many labs in order to achieve a better meta-analytic estimate of the effect size.

Sample demographics:

• Undergraduate subject pool

• Subjects should be students under the age of 35

• Between 20% and 80% female

• Subjects will not be asked to report their ethnic background

• Subjects should be naïve. Labs may use introductory psychology subject pools but they should keep in mind that these pools might be suboptimal as a source of participants for several reasons: First, psychology classes often teach about deceptive experiments, and the effect might be reduced if participants think the experiment might be deceptive. Moreover, experience with studies of this sort as well as with studies in general might moderate the effect. To minimize these concerns, labs should make an effort to collect their data at the start of the semester to limit the number of subject-pool studies completed before this one. Alternatively, laboratories could recruit from a broader campus population with less experience participating in psychology studies. If possible, laboratories should obtain data on actual prior study participation by each participant from their subject pool coordinator. (It is important that potential participants not be asked directly about previous participation as a criterion for inclusion, as doing so would create an incentive for them to under-report their prior experience.)

In order to permit tests of these potential moderators across laboratories, the protocol also will collect self-reports about prior experience with subject pool, paid, or

online experiments in general as well as with studies involving choices about allocating money in particular. Although this study does not involve deception, a further question will ask whether participants had taken part in any deceptive studies before this one as a way to determine whether they might expect deception.

Testing location:
• In-person, laboratory-based testing.
• Subjects must not be able to see the displays or responses of other participants, and they should be confident that no one (including the experimenters) can see their own screen.
• Subjects must be tested in groups of 12 or more people to ensure: (a) that they believe the explanation that the payoff depends on other people, and (b) that they cannot determine which of the other people in the room are in their group.
• The total number of participants in a session must be a multiple of 4. If the number of people who arrive for a testing session is not a multiple of 4, the additional participants should either be paid the show-up fee and asked to participate in a future study or they may complete another unrelated experiment in your lab (provided that additional experiment does not disrupt the participants in the RRR study). For example, if you schedule 16 participants and 1 participant does not show up, you would need to test 12 participants in the RRR study, and the 3 additional participants could take part in an unrelated experiment or receive their show-up fee and leave.
• Note: If your testing room has only 12 computer stations, you schedule 12 participants, and 1 or more do not show up, you can test a group of 8 participants if you have no other option. If your testing room has only 12 stations, and you can schedule 13 or 14 participants to increase your chances of testing groups of 12, we would recommend that you do so.
• When a subject starts the experimental procedure on his/her computer, he/she will be randomly assigned to one of the two conditions. From that point forward, the procedure is automated.

Experimenters:

• Any trained research assistant, postdoctoral researcher, or faculty person.

• Experimenter should have experience collecting experimental psychology data and interacting with small groups of subjects.

• No special expertise is necessary to conduct the study.

• The experimenter will be blind to condition assignment.

Data collection:

• All tasks, questions, and surveys will be administered on a computer using the provided experimental script.

• Subjects should be blind to the hypothesis about the intuitive-cooperation effect and should be unaware of any conditions other than their own (i.e., subjects in the time-pressure condition should not know that there is a forced-delay condition, and vice versa).

• Subjects should be randomly assigned to the experimental and control condition with the constraint that approximately equal numbers of subjects are assigned to each condition. The program script will randomly assign participants to the conditions.

• The signup for the study should describe it as a "study of decision making." No further information should be provided about the content of the study except for its duration, location, and compensation.

Data from a subject may be excluded if:

• They did not complete all tasks.

• The Experimenter/computer incorrectly administered the task or instructions.

• All exclusion decisions must be made by someone who is blind to condition assignment and before examining the amounts of money allocated in the task. Note that all data, including those from excluded participants, must be reported in the Excel template that will be provided to labs. The template will have a column to indicate that data from a particular subject was excluded and an adjacent column to indicate the reason for exclusion.

Procedure and Task Sequence:

1) Subjects are recruited to participate in a study on decision making.

2) The Experimenter opens the Qualtrics survey on each computer in the multi-computer room/lab.

3) Subjects enter the multi-computer room in a group consisting of at least 12 participants. The number of participants tested in each session must be a multiple of 4 (e.g., 12, 16, 20).

4) The participant provides informed consent.

5) The Experimenter seats each participant at a computer.

6) Once all participants are seated, the Experimenter instructs the participants to begin the Qualtrics survey on their computers.

7) The Qualtrics survey randomly assigns that participant to the time-pressure condition or the forced-delay condition.

8) The experiment from that point forward is automated and appears as a series of screens. The procedure below is for USA labs. Laboratories conducting their testing in other languages will need to translate these materials and then have another translator independently translate them back to English to verify their accuracy. The editor will work with these labs to verify the accuracy of their instructions and procedures. Notes on the procedure are presented in Italics and bracketed by "[ ]" below.

Screen 1:
In this task, you will participate in a simple decision making study. You will receive a $5 show-up fee/course credits [*Labs using subject pools might not provide the show-up fee*], and then earn additional money based on your decision and the decision of others. You will be paid in cash immediately following the experiment.

Screen 2:
You have been randomly assigned to interact with 3 of the other people in the room. All of you receive this same set of instructions. You cannot participate in this study more than once.

Each person in your group is given $4 for this interaction.

You each decide how much of your $4 to keep for yourself, and how much (if any) to contribute to the group's common project (from 0 to 400 cents).

All money contributed to the common project is doubled, and then split evenly among the 4 group members.

Thus, for every 2 cents contributed to the common project, each group member receives 1 cent.

If everyone contributes all of their $4, everyone's money will double: each of you will earn $8. But if everyone else contributes their $4, while you keep your $4, you will earn $10, while the others will earn only $6. That is because for every 2 cents you contribute, you get only 1 cent back. Thus you personally lose money on contributing.

The other people really will make this decision too – there is no deception in this study.

Once you and the other people have chosen how much to contribute, the interaction is over. None of you can affect each other's payoffs other than through the single decision in this interaction.

Screen 3:
[Time pressure condition] Please make your decision as quickly as possible. You must **make your decision in less than 10 seconds**!

Please use the slider to choose the amount of money you wish to contribute.

| 0 | 80 | 160 | 240 | 320 | 400 |

Your contribution:

Time remaining: <Timer that counts down from 10>

[Forced delay condition] Please carefully consider you decision. You must **wait and think for at least 10 seconds** before making your decision!

Please use the slider to choose the amount of money you wish to contribute.

| 0 | 80 | 160 | 240 | 320 | 400 |

Your contribution:

Time elapsed so far: <Timer that counts up from 0>

Screen 4 through 6:
- What level of contribution earns the highest payoff for the group as a whole?

- What level of contribution earns the highest payoff for you personally?

[*These questions provide a check on the participant's comprehension of the payoffs.*]

- Please describe why you chose to contribute the amount that you did in the study.

Screens 7 through 27:
These screens will contain the items of the individualism/collectivism scale. The items will be randomly presented to each participant.

Screens 28 through 32:
Experience 1
"To what extent have you participated in studies like this one before? (i.e. where you choose how much to keep for yourself versus contributing to benefit others)".

[*Subjects have to answer on a 5 point Likert scale with 1= Nothing like this scenario, 3 = Somewhat like this scenario, and 5 = Exactly this scenario.*]

Experience 2
"How many total experiments have you completed in exchange for credit or extra credit in your courses?"

[*Subjects have to answer this question by entering a number in a box with 0 as the lowest response option.*]

Experience 3
"How many total paid experiments have you done?"

[*Subjects have to answer this question by entering a number in a box with 0 as the lowest response option.*]

Experience 4
"Have you done any experiments on Mechanical Turk or other online sites? If yes, how many experiments did you do in total online?"

[*Subjects first have to select yes or no in response to the first part of the question. If they select yes, they have to answer the second part of the question by entering a number in a box.*]

Experience 5
"Although this study did not involve deception, have you previously participated in any studies in which you were deceived about the actual purpose of the study? If yes, estimate how many studies you have done that involved deception."

[*Subjects first have to select yes or no in response to the first part of the question. If they select yes, they have to answer the second part of the question by entering a number in a box.*]

Screens 33 through 35:

Trust-in-others
"To what extent do you feel you can trust strangers?"
Subjects have to answer on a ten-point Likert scale from "1=Very Little" to "10=Very Much"

"To what extent do you feel you can trust other people that you interact with in your daily life?"
Subjects have to answer on a ten-point Likert scale from "1=Very Little" to "10=Very Much"

Trust 1
"Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?"

[*Subjects respond by clicking radio buttons for "Most people can be trusted," "Can't be too careful," or "Depends"*]

Trust 2
"Do you think most people would take advantage of you if they got a chance, or would they try to be fair?"

[*Subjects respond by clicking radio buttons for "Would take advantage," "Would try to be fair," or "Depends"*]

Trust 3
"Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?"

[*Subjects respond by clicking radio buttons for "Try to be helpful," "Just look out for themselves," or "Depends"*]

Trust 4
"You can't count on strangers anymore."

[*Subjects respond by clicking radio buttons for "More or less agree" or "More or less disagree"*]

Screens 36 to 39:
These screens will contain the four PARH items

Screen 40:
What year were you born?

Screen 41:
What is your gender?

Screen 42:
In which city and country did you grow up? If you lived in more than one place, list the one you spent the most time in. (If the United States, please also give the state).

Screen 43:
How many other participants in the room do you know?

Note for non-USA labs

The money subjects have available for the one-shot public goods game will be 4 dollars for labs in the USA. However, for labs from outside the USA, it is necessary to adjust this amount to account for differences in currency and the utility of the amount across countries. An appropriate benchmark may be to equalize the amount across countries on the basis of local purchasing power (see http://www.numbeo.com/cost-of-living/compare_countries.jsp). Labs must justify their choice on their OSF page for their replication study, and the editor must approve the choice before data collection begins.

Required Data Analyses:

• The proposing lab (i.e., the Erasmus University Lab) will provide scripts (R scripts or equivalent) to perform all of the analyses required of individual laboratories. All labs must use the provided scripts to conduct their analyses, and we will provide instructions on how to do so. Following completion of the RRR, all scripts used for the meta-analysis and for the moderator analyses will be posted on the Open Science Framework project page.

• <u>Primary analysis</u>: The primary effect size measure is a comparison of the mean percentage contributed in the time-pressure condition and in the forced-delay condition. Participants will be included regardless of whether or not they obeyed the time

constraint, in order to avoid selection effects (see Tinghog et al. 2013 for further discussion), but if more than 10% of the data are excluded for this reason, we will also report the effect size with exclusions. The data will be analyzed both with and without excluding experienced participants who answered anything other than "1" to the following experience question:

"To what extent have you participated in studies like this one before? (i.e. where you choose how much to keep for yourself versus contributing to benefit others)."

The critical effect size is based on the percentage of the money that participants contributed to the public good. Specifically, for each participant, the amount they contributed is divided by the total possible contribution, yielding the proportion contributed (which is then multiplied by 100 to convert to a percentage). Each lab will report the average percentage in the time pressure condition, the average percentage in the time delay condition, and the difference between the two percentages. The central result in the RRR will be a meta-analysis of these percentage differences across experiments.

If there are large differences in the mean percentage contributions across labs (e.g., some labs report 10-20% contributions and others report 70-80% contributions), we will also conduct a meta-analysis using a risk ratio rather than a percentage difference effect size measure and will note if the outcomes of these analyses differ. However, we expect that the contributions in most studies will be in the same range, so the percentage difference will be the most easily interpreted, and hence the preferred, effect size measure.

These effect size estimates will be subjected to additional secondary analyses of moderators (see below). All of these meta-analyses will be reported on the OSF project page, and some may be included in the printed report as well.

• Secondary analyses:
  - To check the random assignment, we will report separately for each condition and lab, the gender composition, mean age, mean comprehension, mean trust in

others, mean experience (as measured by the four separate measures of experience), and mean score on the individualism/collectivism scale.

• <u>Moderator analyses:</u>

- The primary moderator analyses will explore: 1) whether experience affects the size of the main effect, 2) whether trust affects the size of the main effect while including all participants, and 3) whether trust affects the size of the main effect while excluding experienced participants. Exploratory moderator analyses will include all participants and will examine whether the main effect is affected by age, gender, comprehension, individualism/collectivism, whether or not they knew other participants, total studies participated in previously, and participation in deceptive studies.

- Separately for each condition in each lab, we will correlate each participant's contribution with each moderator. These correlations will be reported in a table similar to the one for the randomization checks. We will report the difference in correlations between the two conditions (by converting to z) and then will meta-analyze those difference scores across labs.

- For each of the continuous moderators, we will also conduct an extreme groups analysis, comparing the upper and lower 3rd of participants on each moderator variable within each lab/condition. For trust, we will do this using top 3rd vs bottom 3rd of trust responses over all data from all labs, rather than within each lab.

References

Cone, J., & Rand, D. G. (2014). Time Pressure Increases Cooperation in Competitively Framed Social Dilemmas. *PLoS ONE, 9*(12), e115756. doi: 10.1371/journal.pone.0115756

Cornelissen, G., Dewitte, S., & Warlop, L. (2011). Are Social Value Orientations Expressed Automatically? Decision Making in the Dictator Game. *Personality and Social Psychology Bulletin, 37*(8), 1080-1090.

De Dreu, C. K. W., Dussel, B., & Ten Velden, F. S. (2015). In Intergroup Conflict, Self-sacrifice is Stronger among Pro-social Individuals and Parochial Altruism Emerges specially among Cognitively Taxed Individuals. *Frontiers in psychology, 6.* doi: 10.3389/fpsyg.2015.00572

Døssing, F., Piovesan, M., & Wengstrom, E. (In prep). Cognitive Load and Cooperation.

Dreber, A., Fudenberg, D., Levine, D. K., & Rand, D. G. (2014). Altruism and Self-Control. *Available at SSRN: http://ssrn.com/abstract=2477454*.

Lotz, S. (2014). Spontaneous Giving Under Structural Inequality: Intuition Promotes Cooperation in Asymmetric Social Dilemmas. *Available at SSRN: http://ssrn.com/abstract=2513498*.

Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature, 489*(7416), 427-430.

Rand, D. G., Greene, J. D., & Nowak, M. A. (2013). Reply to "Intuition and cooperation reconsidered". *Nature, 497*(7452), E3.

Rand, D. G., & Kraft-Todd, G. T. (2014). Reflection Does Not Undermine Self-Interested Prosociality. *Frontiers in Behavioral Neuroscience, 8*, 300.

Rand, D. G., Newman, G. E., & Wurzbacher, O. (2015). Social context and the dynamics of cooperative choice. *Journal of Behavioral decision making, 28*, 159-166.

Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., & Green, J. D. (2014). Social Heuristics Shape Intuitive Cooperation. *Nature Communications, 5*, 3677.

Roch, S. G., Lane, J. A. S., Samuelson, C. D., Allison, S. T., & Dent, J. L. (2000). Cognitive Load and the Equality Heuristic: A Two-Stage Model of Resource Overconsumption in Small Groups. *Organizational Behavior and Human Decision Processes, 83*(2), 185-212. doi: 10.1006/obhd.2000.2915

Schulz, J. F., Fischbacher, U., Thöni, C., & Utikal, V. (2014). Affect and fairness: Dictator games under cognitive load. *Journal of Economic Psychology, 41*, 77-87. doi: http://dx.doi.org/10.1016/j.joep.2012.08.007

Singelis, T. M., Triandis, H. C., Bhawuk, D. P., & Gelfand, M. J. (1995). Horizontal and vertical dimensions of individualism and collectivism: A theoretical and measurement refinement. *Cross-cultural research, 29*(3), 240-275.

Tinghög, G., Andersson, D., Bonn, C., Böttiger, H., Josephson, C., Lundgren, G., . . . Johannesson, M. (2013). Intuition and cooperation reconsidered. *Nature, 497*(7452), E1-E2.

Verkoeijen, P. P. J. L., & Bouwmeester, S. (2014). Does Intuition Cause Cooperation? *PLoS ONE, 9*(5), e96654. doi: 10.1371/journal.pone.0096654