

IZA DP No. 9871

Intuitive Cooperation and Punishment in the Field

Luis Artavia-Mora
Arjun Bedi
Matthias Rieger

April 2016

Intuitive Cooperation and Punishment in the Field

Luis Artavia-Mora

ISS, Erasmus University Rotterdam

Arjun Bedi

*ISS, Erasmus University Rotterdam
and IZA*

Matthias Rieger

ISS, Erasmus University Rotterdam

Discussion Paper No. 9871
April 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Intuitive Cooperation and Punishment in the Field*

We test whether humans are intuitively inclined to cooperate with or punish strangers using a natural field experiment. We exogenously vary the time available to help a stranger in an everyday situation. Our findings suggest that subjects intuitively tend to help but behave more selfishly as thinking time increases. We also present suggestive evidence that time pressure can increase rates of punishment. We discuss our results with respect to findings in the lab on cognitive models of dual-processing and the origins of human cooperation.

JEL Classification: D03, D63, D64

Keywords: cooperation, punishment, response time, dual-process of cognition, natural field experiment

Corresponding author:

Luis Artavia-Mora
International Institute of Social Studies (ISS)
Erasmus University Rotterdam
Kortenaerkade 12
2518 AX Den Haag
The Netherlands
E-mail: artaviamora@iss.nl

* We thank the Economics of Development and Emerging Markets group and Economics of Development study program at ISS for funding the data collection. We received valuable comments from Brandon Restrepo, Brigitte Vézina and Maria Dafnomili. This paper is based on and a substantially reworked version of the MA thesis by Luis Artavia-Mora entitled "Intuitive cooperation in The Hague: A natural field experiment" (ISS Working Paper 614) supervised by Matthias Rieger and Arjun Bedi.

1 INTRODUCTION

This paper tests whether humans are intuitively inclined to cooperate and punish in one-shot interactions in the field. Specifically, we investigate how time pressure and time delay impact the likelihood of helping or punishing strangers in real life. To this end we propose a novel, natural field experiment based on a “dual-process cognitive framework.” This framework contrasts intuitive versus deliberate decision-making to understand the origins of human cooperation (Rand et al. 2012, 2014). Thinking intuitively refers to faster and automatic decisions based on prior experience, beliefs and instinct, while thinking in a deliberate way refers to slower, controlled and more reflective decisions (see Rand et al. 2012, 2014; as well as, Loewenstein and O’Donoghue, 2004; Frankish, 2009; Kahneman, 2012; Evans and Stanovich, 2013).

We build on previous findings that humans are naturally predisposed towards cooperation, but tend to behave more selfishly as thinking time increases. Time pressure has a positive impact on rates of cooperation in the lab (Rand et al, 2012; Rand et al. 2014). Similarly, people that contribute in a public good game tend to make faster decisions than free riders (Nielsen et al, 2014). Conversely, other studies have found that faster responses are linked to more egoistic decisions (Piovesan and Wengström, 2009), while a third group of studies has not found clear distinctions (Tinghög et al, 2013; Verkoeijen and Bouwmeester, 2014). In response to these somewhat mixed empirical findings, Rand et al. (2014, p.2) proposed the *social heuristics hypothesis* which posits that “daily life typically involves factors such as repetition, reputation and the threat of sanctions, all of which can make cooperation in one’s long term self-interest.” This in turn generates “generalized cooperative intuitions.” Put differently, the theory directly links learning from experience and daily interactions with behavioral outcomes. Personal experiences with social norms could ultimately shape selfish or cooperative predispositions. And these everyday experiences may or may not lead to intuitive cooperation in the lab.

We make three contributions to the literature on the origins and drivers of human cooperation: First, we examine the origins of human cooperation in a natural field setting. We assess the drivers of cooperative behavior in a realistic way (List, 2007; List, 2011). Subjects will draw on their everyday experiences when making their decisions.

Second, we conjecture that our setting minimizes the possibility of conflating human mistakes and confusion under time pressure, which is something that may plague laboratory experiments (Recalde et al., 2014). Rather than pushing buttons, subjects will provide actual help to a stranger at a personal cost (such as bending down and picking up a dropped object).

Third, we are—to the best of our knowledge—the first to test the impact of response time on both cooperation and punishment. We test if response time impacts the likelihood of punishing or cooperating with norm violators. If response time has pro-cooperation effects, then examining intuitive punishment is an interesting cross-validation exercise. Punishing a stranger who violates a social norm may benefit society, but also comes at a personal cost and involves fear of retaliation. Our hypothesis is that longer thinking time decreases impulsive actions. If direct punishment is indeed impulsive, it should be less frequent under time delay. Direct punishment rates in the field are typically low; so we also examine indirect punishment, which we define as withholding help from a norm violator (see Balafoutas et al. 2012; Balafoutas et al. 2014). Unlike direct punishment withholding help is a more thought-through form of punishment. So our hypothesis is that indirect punishment is less likely to be affected by time pressure.

Our experiment uses actors in the field to trigger opportunities for subjects to help and punish (Balafoutas et al. 2012; Balafoutas et al. 2014). We propose a new methodology to randomly vary response times by manipulating distances between actors and subjects. In our experiment subjects have either about 3.5 seconds or 10 seconds to make their decision to cooperate or not. The basic aim of such exercises is to get a sense of the “dominant directions of the effects of intuitive versus reflective processing ” (see Rand et al. 2014, p.2); and of course “[re]flection may fail to override deeply held intuitions, and some subjects may engage in substantial reflection even under time pressure.”

We document that cooperation rates decline from 71% to 52% when subjects have more time to think. We take the direction of the effect as suggestive evidence that - on average - humans are intuitive cooperators. Selfishness rises as response time increases. Similar patterns occur in the case of direct punishment. We present suggestive evidence that time delay roughly halves rates of direct punishment of norm violators. Indirect punishment is not significantly affected by time

pressure. We also investigate heterogeneities and mechanisms underlying these average impacts by actor and subject characteristics. Experience with local norms of cooperation as proxied by the years lived in the study country, as well as risk preferences are only weakly correlated with cooperation rates. And there is some evidence that the positive impact of time pressure on the likelihood of helping a stranger is concentrated among risk averse individuals. The impact of response time is statistically smaller among risk taking individuals. In sum, our results are in line with studies showing a positive link between time pressure and pro-social behavior (Rand et al, 2012; Rand et al. 2014).

Our paper is organized as follows. Section 2 details the experimental design and data. Section 3 presents the results. Section 4 compares findings to related studies and concludes.

2 THE EXPERIMENT

Our hypothesis is that time pressure increases cooperation and punishment. This section details the two experiments to test this hypothesis along with the choice of field location, treatments, as well as the experimental procedures and subject characteristics.

2.1 TWO SOCIAL DILEMMAS

We designed an experiment featuring two social dilemmas. In the *first* dilemma we triggered help from subjects by asking an actor to drop one bicycle glove in a public park (as shown in Photos 1 and 2 in the Annex). We interpret the decision to help or defect as the choice between cooperation at a personal cost versus selfishness. We chose the glove drop for four simple reasons: First, gloves are complements. Losing one glove makes the second glove useless. Second, a glove falls noiselessly and it is thus credible that the actor does not notice the loss of the glove and requires help. Third, gloves are big enough to be seen from a distance which is necessary for our response time treatments as we explain below. Fourth, a glove is neither too cheap nor too expensive. We wanted to minimize distortions to cooperate or defect based on the value of the object. Using an expensive object (i.e. jewelry) may lead to higher rates of help but also potentially to theft. Conversely, using a cheap object (i.e. pencil) may dissuade help and might be perceived as littering. Arguably, gloves are a good compromise.

The *second* dilemma extends the glove experiment by adding the violation of a social norm. Specifically, we asked the actor to litter just before dropping the glove. The actor litters (throws an empty plastic bottle) and then drops the glove (see Photo 3 in the Annex). The idea is to investigate whether humans punish directly (verbal punishment) or indirectly (withholding help) if an individual litters and to assess the impact of response time on punishment.

2.2 LOCATION

The location of the experiment was a pedestrian path in Park Malieveld in The Hague in the Netherlands. The location is appropriate for at least three reasons: First, the path is straight and bordered by trees. It is the only way to cross the park. Therefore it is hard for subjects to avoid or dodge the decision to help or not (see Photos 1 and 5 in the Annex). More importantly, there is little distraction and visibility is good (see Photos 6 and 9 in the Annex). It is easy to see the glove drop. Second, based on prior observation we noticed that people on the path tend to walk alone and that there can be large distances between them. This is important since we only wanted to sample subjects walking alone. Confounds such as reciprocity and social pressure are thus minimized. Subjects can make “private” and “anonymous” decisions in a public space (Photos 1 to 9 in the Annex). Third, the location is near the heart of the city and is surrounded by many stores, institutions and workplaces (i.e. government offices, learning institutes, shopping areas, university faculties, commercial businesses and non-governmental organizations). This yields a relatively diverse pool of subjects.

2.3 EXPERIMENTAL TREATMENTS

Our experiment tests the impact of response time on cooperation with strangers. We generated exogenous variation in response time by varying the distance between subject and actor. Average human walking time is about 1.3 meters per second (Mohler et al. 2007). We use two distance treatments, one short and one long. The short distance is 4.5 meters between the subject and actor. This provides roughly 3.5 seconds for an individual to decide whether to help or defect. This treatment elicits decisions under time pressure. The longer distance is 13 meters and subjects have 10 seconds to decide. The longer time period is designed to elicit a deliberate decision. We calibrated distances based on visibility. If the glove is dropped at a distance closer than 4.5 meters, the field of vision is too narrow and restricted. If the glove is dropped from a distance further than 13 meters, visibility declines.

Figure 1 summarizes the two distance treatments. The actor is depicted in grey and the subject in black. Point A indicates the location where the actor drops a glove and triggers the social dilemma. The actor randomly drops the glove either when the subject is at point B (4.5 meters) or at point C (13 meters). Due to the random assignment of subjects to points, we can isolate the effect of distance on the likelihood of helping the actor.

In the punishment extension, the actor litters (violates the non-littering norm) **before** the participant reaches points B (or C). Then the actor drops a glove at point B (or C). Table 1 summarizes the 2x2 design of the natural field experiment.

2.4 SUBJECT SELECTION AND EXPERIMENTAL PROCEDURE

Unless weather conditions were not suitable (rain, storm), the field experiment was performed during 11 days in July 2015. The survey took place on working days between 10:00 am and 5:00 pm. The treatments were randomly assigned and are thus independent of subject type, weekday and time of the day. We used one female and one male actor at random. A researcher recorded the data and was located at a distance to avoid social pressures (see Photo 2 In the Annex).

Each trial began when the researcher selected a participant. The selection was based on two criteria: First, the subject needed to be alone with no other individual walking in the same or opposite direction. This criterion was imposed to eliminate social pressures. Second, the participant had to be in no visible hurry nor visibly distracted. Photos 4 and 5 in the Annex show a typical participant in the experiment. In a few cases subjects were not selected since the actor or surveyor knew the subject personally.

Each experiment started with the actor sitting on a bench at Point A. Parked bicycles marked points A, B and C (see Figure 1). The actor held a pair of gloves (and if applicable the plastic bottle for the social norm violation scenario) and carried a bag (see Photos 1 and 2 in the Annex). The actor then left the bench and started crossing the path, waiting for the participant to reach point B or point C. When the subject reached either point B or C, the actor “accidentally” dropped one bike

glove without noticing while trying to put it in a bag. The actor then stopped at his/her bicycle and pretended to be looking for the keys of the bicycle or to make a phone call (see Photos 1 and 3 in the Annex). The actor waited until the participant revealed the decision to cooperate (or defect) at point A.

It is important to note that the actor ignored any voice alerts far from point A. Instead, the actor only responded at point A (see Photos 10 and 12 in the Annex). This ensured that each participant had the same time to help the actor.

In the last step of the experiment, that is once the participant had made a decision at point A, the researcher noted down the results, while the actor quickly interviewed the participant (Photos 13 and 14 in the Annex). The survey collected demographic characteristics such as age, gender, time lived in The Netherlands, willingness to undertake risks in daily life and height as a proxy of physical strength and confidence (the short survey can be found in the Annex). We use these variables to investigate treatment balance, as well as impact heterogeneity.

To summarize, each participant's response time depended on individual walking speed, but most importantly on the randomized distance to the dropped glove. The treatment with the shorter response time was designed to elicit intuitive decisions, while the time delayed condition was meant to promote deliberate decisions.

2.5 SUBJECT CHARACTERISTICS AND TREATMENT BALANCE

We ran 267 trials – 137 for helping a stranger and 130 for helping the norm violator. Table 2 shows that subject characteristics are balanced across treatments suggesting that randomization was achieved. These basic demographic statistics stem from the post-experiment survey.¹ The

¹ Note that the response rate to the survey was 88% and non-response is unrelated to the distance and social dilemma treatment at the 5% level of significance (see p-values in Table 2). There is a somewhat lower response rate in the case of the littering experiment, which could be interpreted as a form of punishment of the actor (6.6 percentage point difference in means; p-value=0.096). In any case, we show that impacts are stable for the whole sample and the smaller sample of people who responded to the survey.

participants have an average age of 44 years and 61% are male. The average subject has lived for 37 years in the Netherlands. 76.5% of people in our sample have lived their entire lives in the Netherlands. These two variables may proxy experience of interacting with strangers and may affect the behavior of subjects.

While we picked distances to ensure maximum visibility, people might not have seen the drop of the glove. This is not a problem *per se* if visibility issues are not systematically related to the distance treatment. After the experiment we asked people if they had noticed the glove drop (see post-experimental survey in the Annex). 93% of subjects acknowledged seeing it. There are no significant differences in this variable between the time pressure and time delay treatments (see p-values in Table 2). In a robustness check below, we show that excluding the people that did not see the glove drop from the analysis yields qualitatively similar results. Therefore our main models includes these outliers as non-cooperators.

3 RESULTS

Response time impacts cooperation rates in our experiment. We also present suggestive evidence on the effects of time delay on punishment. The overall evidence indicates that humans are naturally inclined to cooperate but behave more selfishly when they have more time to think. Basic cooperation results are summarized in Figure 2. Panel A shows rates of helping a stranger (who dropped a glove) by response time treatment. Panel B gives the corresponding rates of helping a norm violator (i.e. littering plus drop of a glove). Figure 3 plots rates of direct punishment of the norm violator.

Helping a stranger

Differences in mean cooperation rates are sizeable (see Panel A, Figure 2). Under time pressure, 71% of participants help the actor who has lost a glove. This percentage drops to 52% when subjects have more time to respond. The 19 percentage point treatment effect underlines that participants are substantially more predisposed to cooperate when the time available to think is short.

Helping a norm violator

Similar patterns emerge in the experiment on helping the norm violator (see Panel B, Figure 2). It is not surprising that overall rates of helping drop when the actor litters before losing a glove. The reduction amounts to 17 percentage points in the case of time delay and 23 percentage points in the case of time pressure (compare panels A and B by treatment group in Figure 2). These uniform overall reductions due to the littering treatment are statistically significant with p-values below 0.05. In other words, the initial littering causes subjects to significantly reduce help, that is to say, punish indirectly. The impact of response time is in line with the previous scenario. There is evidence that subjects are more likely to defect as response time increases. The difference in means amounts to 13.6 percentage points but the estimate is imprecise with a p-value of 11.8%.

Comparing the two social dilemmas

Time delay impacts are comparable across social dilemmas. More time to think causes 27% (helping a stranger) and 28% (helping the norm violator) reductions in helping rates, respectively. Table 3 allows a direct comparison between the two experiments using a regression model. Pooling the two experiments also increases the efficiency of the estimates. Both time delay and the violation of the social norm decrease helping rates (see columns 1 and 2). The two treatment effects are similar in magnitude and statistically significant (column 3). Does time delay magnify or decrease the effect associated with norm violation? To answer this question note that the interaction between the two treatments (time delay and norm violation) is positive but insignificant (column 4). In other words, there is weak evidence that time delay reduces the negative effect of norm violation on the likelihood of helping a stranger (in absolute terms).

Punishment

Impacts of response time on *direct* punishment of littering are presented in Figure 3. Overall 13.85% of subjects punish directly (by voice). While 18.75% of subjects directly punished the actor in the case of time pressure this rate halves to 9.09% when thinking time increases. While the treatment effect is large it is marginally insignificant with a p-value of 0.113. Note that the direct punishment rate under time delay is similar to that found in previous studies. Balafoutas et al. (2012) and Balafoutas et al. (2014) report rates as low as 4% and 6.8% in littering experiments

in train stations in Greece and Germany, respectively. Time delay possibly offers one explanation for such low rates. We can also look at the interaction between helping and direct punishment (see Table 4). 10 out of 130 people punished directly but still helped.

How does time delay impact the form of punishment? 76 out of 130 people did not help the norm violator. Out of these, 8 punished directly. Cell counts are small, but there is suggestive evidence that time pressure favors direct over indirect punishment. All of these 8 direct punishers acted under time pressure. Under time delay none of these 76 subjects punished directly.

Robustness and inclusion of co-variables

Table 5 presents regression models for the three binary response variables (helping, helping the norm violator, direct punishment of the violator) where we include subject covariates. Due to the randomization of treatments, point estimates associated with time delay are stable across models. Note also that covariates have relatively little explanatory power. While risk taking individuals, as well as those who have lived their entire life in the Netherlands (so-called natives) are more likely to help and punish, the effects are imprecisely estimated.² Finally, we ran an unreported robustness check including the time of the survey (morning vs. afternoon) and a dummy for the day of the week as co-variables. The coefficients associated with the time delay treatment is stable, as also indicated by tests of the equality of coefficients using seemingly unrelated regression models.

Treatment heterogeneity

Table A2 in the Annex examines stability of estimates across sub-samples of actors as well as subject characteristics. First, there could be a concern that subjects did not notice the actor or the glove drop itself. Row 1 splits the sample into people that acknowledged witnessing the drop (93%)

² We have coded behavior as a binary outcome: help or not. We also examined the various behaviors prior to and when cooperating. We classified a full range of responses in Table A1 in the Appendix. Counts of helping behaviors are given conditional on helping. Typical helping behaviors are illustrated in Photos 10 and 11 in the Annex. When it comes to behavior leading up to the decision, the majority of people showed no reaction, followed by voice alert. There are no systematic differences between time pressure and time delay treatments.

versus those that did not. Differences between point estimates are small and statistically insignificant. Second, we used two actors – one male and one female. One could be worried that the gender of the actor influences behavior, which may be problematic for the external validity of our results. Row 2 indicates that our point estimates are stable across actors. Of course, more actors are needed to investigate actor-specific traits such as height or ethnicity that may influence behavior. Third, subject characteristics matter little - with one exception (see Rows 3-7). The impact of time delay is smaller for risk-taking individuals (the difference is significant at the 10% level). However, this pattern does not hold for helping the norm violator. In sum, we can document little consistent heterogeneity in the treatment impacts, although we may be lacking power for such a finely grained exercise.

4 DISCUSSION AND CONCLUSION

We contribute to the literature on the origins of human cooperation using an original natural field experiment. We document pro-cooperation effects of time pressure in line with lab experiments (see Rand et al. 2012; Rand et al. 2014; Neilsen et al. 2014). In addition, we provide suggestive evidence that time pressure increases rates of direct punishment.

Similar to previous studies and on the basis of our findings we argue that, on average, humans are naturally predisposed to help strangers. To examine this more explicitly we asked our subjects if they found it difficult to make the decision to help or not (see post-experimental survey in the Annex). The responses indicate that time delay made it significantly harder to make a decision. The proportion of people that reported “quite a lot” or “a lot” of difficulty was 26 percentage points higher for those in the time-delayed treatment in the case of helping a stranger (p -value=0.001) and 15 percentage points in the case of helping a norm violator (p -value=0.139).

How do rates of cooperation compare with previous studies? Rand et al. (2014) find that under time delay, contributions to public good games decrease by 21%. In our binary set-up, time delay leads to 27% (helping a stranger) and 28% (helping the norm violator) reductions in helping rates. What could be the underlying mechanism behind our results? Rand et al. (2014) indicate that experience in a given place and with the associated social norms play an important role in shaping an individual’s decisions and reflexes to cooperate. Foreigners might behave differently than locals

in our experiments. We proxy experience by comparing subjects who have lived their entire lives in the Netherlands versus others. However, native subjects were not significantly more likely to help and punish in our experiments.

One contribution of this paper is that we can alleviate the concern that one too easily conflates human error and intuition in a lab setting (Recalde et al. 2014). While of course we cannot fully rule out errors or confusion, a field experiment and an everyday situation offers a natural setting. “Bending down and picking up a glove” as many of our subjects did is plausibly less error-prone than pushing a button in the lab. We also documented that the visibility of the glove drop was good and that the vast majority of subjects acknowledged seeing it. More importantly, visibility is unrelated to the distance treatments. We cannot, however, rule out that defectors made an error under time pressure. Fortunately, such errors would work against us finding a pro-social effect of time pressure.

Our study opens avenues for future research: First, participants have different abilities to digest information and make decisions. While randomization ensures balance across ability types, the average effects documented in our paper may conceal heterogeneity in terms of individual processing speeds (Rubinstein, 2007). Second, the impacts of time delay may be specific to the helping task and context. Future work may investigate the stability and universality of our estimates across space, time and task (Rieger and Mata, 2015). The speed, complexity and importance of the dilemma itself may influence human behavior (see related studies by Fehr and Rangel, 2011). Third, it may be worth investigating heterogeneous effects specific to actor traits. Finally, it would be interesting to examine intuitive cooperation among children as they age and engage in new experiences.

References

- Balafoutas, L., and N. Nikiforakis. 2012. "Norm Enforcement in the City: A Natural Field Experiment." *European Economic Review* 56 (8): 1773-1785
- Balafoutas, L., N. Nikiforakis, and B. Rockenbach. 2014. "Direct and Indirect Punishment among Strangers in the Field." *Proceedings of the National Academy of Sciences* 111 (45): 15924-15927.
- Evans, J.S.B.T., and K.E. Stanovich. 2013. "Dual-Process Theories of Higher Cognition: Advancing the Debate." *Perspectives on Psychological Science* 8 (3): 223-241.
- Fehr, E., and A. Rangel. 2011. "Neuroeconomic Foundations of Economic Choice — Recent Advances." *The Journal of Economic Perspectives* 25 (4): 3-30.
- Frankish, K., and J.S.B.T. Evans. 2009. *The Duality of Mind: An Historical Perspective*. Oxford, UK: Oxford University Press.
- Gächter, S. 2012. "A Cooperative Instinct." *Nature* 489 (7416): 374.
- Kahneman, D. 2012. *Thinking, Fast and Slow*. London, UK: Penguin.
- List, J.A. 2007. "Field Experiments: A Bridge between Lab and Naturally Occurring Data." *The B.E. Journal of Economic Analysis and Policy* 6 (2): 8.
- List, J.A. 2011. "Why Economists should Conduct Field Experiments and 14 Tips for Pulling One Off" *The Journal of Economic Perspectives* 25 (3): 3-15.
- Loewenstein, G., and T. O'Donoghue. 2004. "Animal Spirits: Affective and Deliberative Processes in Economic Behavior." Working paper 04-14, Cornell University, Center for Analytic Economics.
- Mohler, B.J., W.B. Thompson, S.H. Creem-Regehr, H.L. Pick Jr., and W.H. Warren Jr. 2007. "Visual Flow Influences Gait Transition Speed and Preferred Walking Speed." *Experimental Brain Research* 181 (2): 221-228.
- Nielsen, U.H., J.R. Tyran, and E. Wengström. 2014. "Second Thoughts on Free Riding." *Economics Letters* 122 (2): 136-139.
- Nowak, M.A. 2006. "Five Rules for the Evolution of Cooperation." *Science* 314 (5805): 1560-1563.
- Piovesan, M., and E. Wengström. 2009. "Fast Or Fair? A Study of Response Times." *Economics Letters* 105 (2): 193-196.
- Rand, D.G., J.D. Greene., and M.A. Nowak. 2013. "Rand Et Al. Reply." *Nature* 498 (7452): E3.
- Rand, D.G., J.D. Greene., and M.A. Nowak. 2012. "Spontaneous Giving and Calculated Greed." *Nature* 489 (7416): 427-430.

Rand, D.G., and M.A. Nowak. 2013. "Human Cooperation." *Trends in cognitive sciences* 17 (8): 413-425.

Rand, D.G., A. Peysakhovich, G.T. Kraft-Todd, G.E. Newman, O. Wurzbacher, M.A Nowak, and J.D. Greene. 2014. "Social Heuristics Shape Intuitive Cooperation." *Nature communications* 5.

Recalde, M. P., A. Riedl, and L. Vesterlund. 2014. "Error Prone Inference from Response Time: The case of Intuitive Generosity." CESifo Working Paper Series No. 4987

Rieger, M., and R. Mata. 2015. "On the Generality of Age Differences in Social and Nonsocial Decision Making." *Journals of Gerontology Series B-Psychological Sciences and Social Sciences* 70 (2): 200-212.

Rubinstein, A. 2007. "Instinctive and Cognitive Reasoning: A Study of Response Times." *The Economic Journal* 117 (523): 1243-1259.

Verkoeijen, P., and S. Bouwmeester. 2014. "Does Intuition Cause Cooperation?" *PloS one*, 9 (5): e96654.

Tinghög, G., D. Andersson, C. Bonn, H. Böttiger, C. Josephson, G. Lundgren, D. Västjäll, M. Kirchler, and M. Johannesson. 2013. "Intuition and Cooperation Reconsidered." *Nature* 498 (7452): E1-E2.

6 FIGURES

Figure 1: Diagram of time pressure and time delay treatments

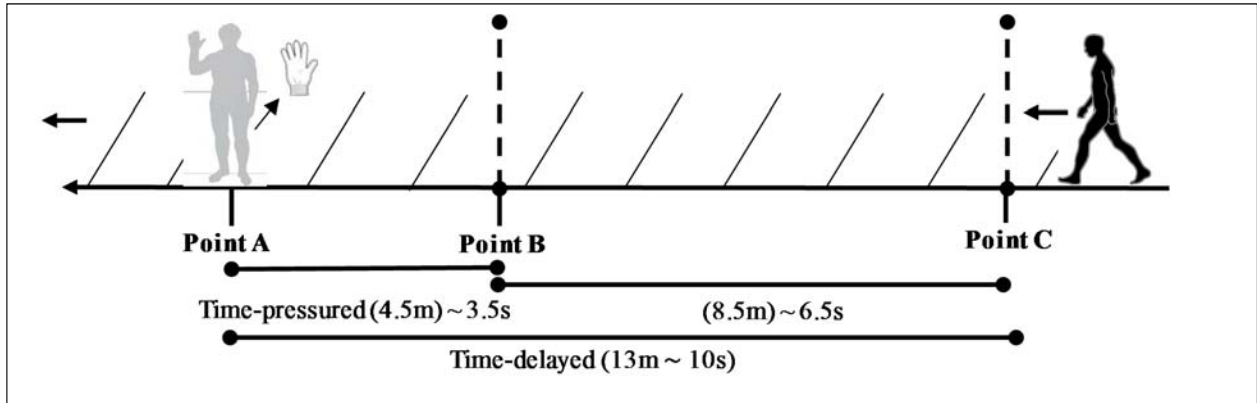
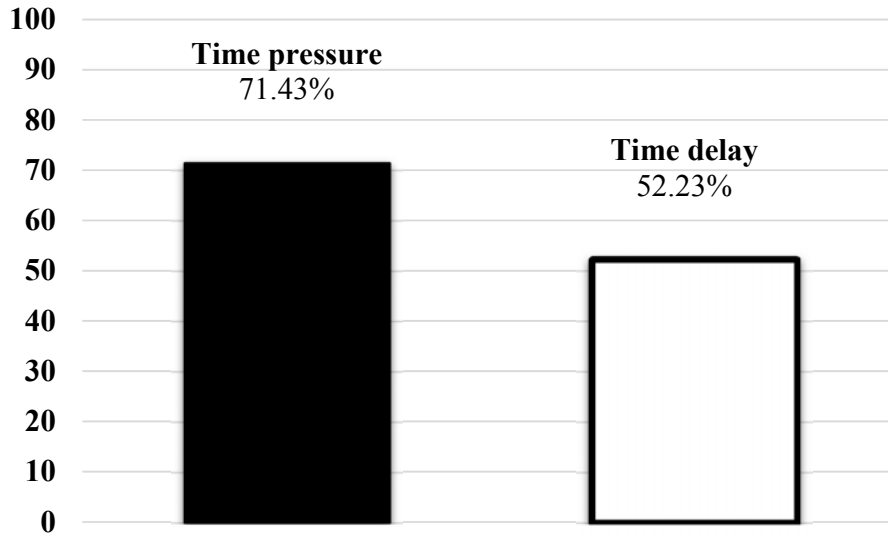


Figure 2: Average rates of cooperation and response time

Panel A: Helping a stranger (Δ means p-value 0.021; n=137)



Panel B: Helping a norm violator (Δ means p-value 0.118; n=130)

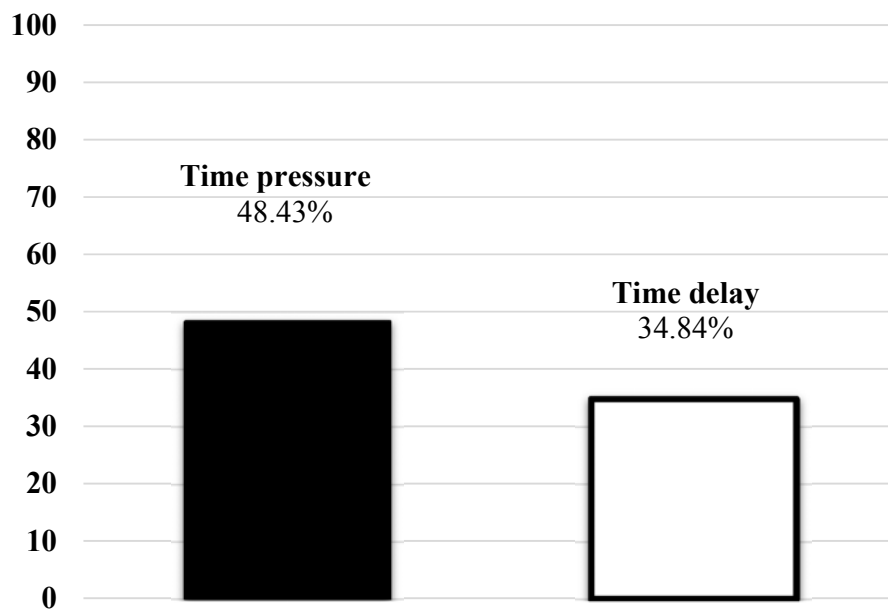
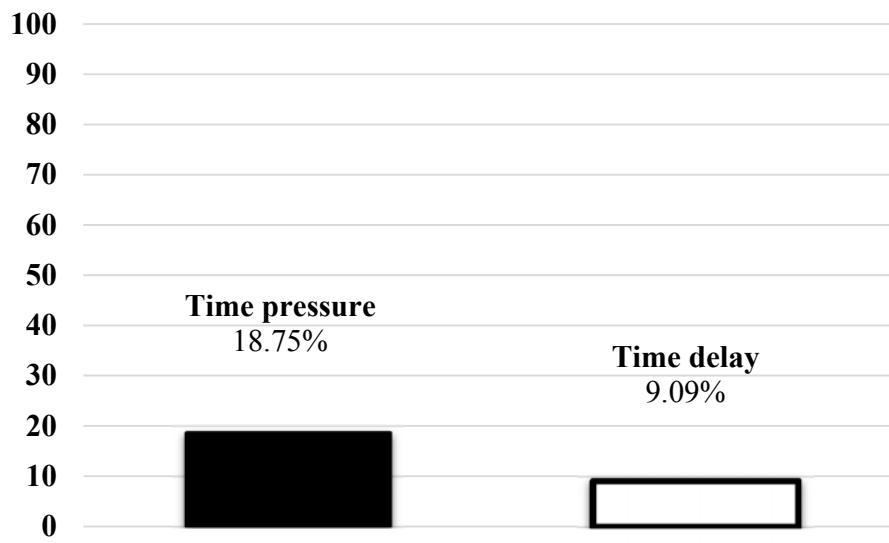


Figure 3: Direct verbal punishment of norm violator (Δ means p-value 0.113; n=130)



8 TABLES

Table 1: Dimensions and treatments in the natural field experiment

Social dilemma	Response time
Helping a stranger <i>(cooperation)</i>	Time pressure <i>(intuition)</i>
Helping a norm violator <i>(cooperation with punishment)</i>	Time delay <i>(deliberation)</i>

Table 2: Characteristics of subjects and balance across treatments (pooled sample)

Variables	N	Mean	SD	Min	Max	Randomization balance	
						<i>Response time</i>	<i>Dilemma type</i>
						P-values Δ	
Responded to survey	267	88.01		0	1	0.98	0.10
Age	234	43.78	13.97	15	76	0.56	0.69
Male (1=male; 0=female)	267	0.61		0	1	0.29	0.87
Height (in cm)	230	175.32	10.77	147	204	0.46	0.23
Years lived in the Netherlands	234	37.74	20.38	0	76	0.70	0.59
Native	234	0.76		0	1	0.31	0.62
Willingness to take risks (0=lowest;10=highest)	233	5.74	1.85	0	10	0.44	0.23
Acknowledged seeing glove drop	267	0.93		0	1	0.47	0.72

Note: Native is defined as having always lived in the Netherlands.

Table 3: Regression results pooling helping (n=137) and helping norm violator experiments (n=130)

Dep. var. Helping	(1)	(2)	(3)	(4)
Time delay	-0.168*		-0.165*	-0.192*
	(0.061)		(0.059)	(0.082)
Norm violator		-0.205*	-0.202*	-0.230*
		(0.060)	(0.059)	(0.083)
Time delay x Norm violator				0.056
				(0.119)
Constant	0.604*	0.620*	0.701*	0.714*
	(0.042)	(0.042)	(0.048)	(0.054)
N	267	267	267	267

Notes: Robust standard errors in parentheses. Symbols denote significance levels at +p<0.1, *p<0.05.

Table 4: Interaction between helping the norm violator and direct punishment (totals)

	Punish	No punish	Total
No help	8	68	76
Help	10	44	54
Total	18	112	130

Table 5: Robustness to inclusion of covariates

Dep. variable	Helping a stranger			Helping a norm violator			Direct verbal punishment of norm violator		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Time delay	-0.192*	-0.183*	-0.213*	-0.136	-0.167+	-0.163	-0.097	-0.119+	-0.103
	(0.082)	(0.085)	(0.083)	(0.086)	(0.098)	(0.100)	(0.061)	(0.071)	(0.068)
<i>Subject characteristics</i>									
Age			0.006			0.002			0.007*
			(0.004)			(0.004)			(0.002)
Male			0.078			-0.215+			-0.085
			(0.115)			(0.115)			(0.070)
Native			0.119			-0.089			0.060
			(0.114)			(0.134)			(0.072)
Risk taking			0.020			0.017			0.029+
			(0.025)			(0.027)			(0.017)
Height (in cm)			0.000			0.004			0.000
			(0.006)			(0.006)			(0.004)
Constant	0.714*	0.742*	0.244	0.484*	0.542*	-0.217	0.188*	0.208*	-0.286
	(0.054)	(0.054)	(0.925)	(0.063)	(0.073)	(1.000)	(0.049)	(0.059)	(0.607)
N	137	125	125	130	104	104	130	104	104

Notes: Robust standard errors in parentheses. Symbols denote significance levels at +p<0.1, *p<0.05.

9 ANNEX

(Not for publication - online supplementary tables, photos and post-experimental survey)

9.1 APPENDIX TABLES

Table A1: Helping behaviors by treatments (in %)

Stage	Behavior	Helping a stranger		Helping a norm violator	
		<i>Time pressure</i>	<i>Time delay</i>	<i>Time pressure</i>	<i>Time delay</i>
<i>Behavior before decision</i>	Looks around	12	5.7	6.5	13
	Hesitates	14	20	29	13
	Voice alert	28	28.6	22.6	17.4
	No reaction	46	45.7	41.9	56.6
<i>Helping behavior</i>	Physical contact & points	42	40	35.5	34.8
	Bends down & picks up glove	32	25.7	22.6	30.4
	Voice alert & points	26	34.3	41.9	34.8

Table A2: Time delay impact heterogeneity in sub-samples

		(1) Helping a stranger	(2) Helping norm violator
(1)	Full sample	-0.192* (0.082)	-0.136 (0.086)
	Acknowledged seeing glove drop	-0.178* (0.0831)	-0.133 (0.090)
	<i>P-value Δ</i>	0.598	0.907
(2)	Actor 1	-0.207+ (0.115)	-0.129 (0.121)
	Actor 2	-0.176 (0.116)	-0.146 (0.122)
	<i>P-value Δ</i>	0.852	0.918
(3)	Male subject	-0.186+ (0.103)	-0.113 (0.107)
	Female subject	-0.190 (0.133)	-0.198 (0.140)
	<i>P-value Δ</i>	0.983	0.630
(4)	Age above median	-0.272* (0.097)	-0.165 (0.128)
	Age below median	-0.161 (0.125)	-0.176 (0.140)
	<i>P-value Δ</i>	0.481	0.954
(5)	Native	-0.208* (0.092)	-0.199+ (0.106)
	Other	-0.219 (0.188)	-0.071 (0.207)
	<i>P-value Δ</i>	0.959	0.584
(6)	Height above median	-0.164 (0.105)	-0.144 (0.139)
	Height below median	-0.215 (0.133)	-0.208 (0.137)
	<i>P-value Δ</i>	0.762	0.746
(7)	Risk taking above median	-0.062 (0.109)	-0.213 (0.131)
	Risk taking below median	-0.375* (0.128)	-0.101 (0.137)
	<i>P-value Δ</i>	0.062	0.554

Note: P-values below estimates stem from tests of the equality of coefficients. We use full sample medians to investigate heterogeneities in terms of subject's age, years lived in the Netherlands, height and risk preferences. Native refers to those who have always lived in the Netherlands. Models on the sub-samples are jointly estimated using seemingly unrelated regressions. Robust standard errors in parentheses. Symbols denote significance levels at +p<0.1, *p<0.05.

9.2 PHOTOS

Photo 1: Helping a stranger in the time pressure treatment³



Photo 2: Helping a stranger in the time delay treatment



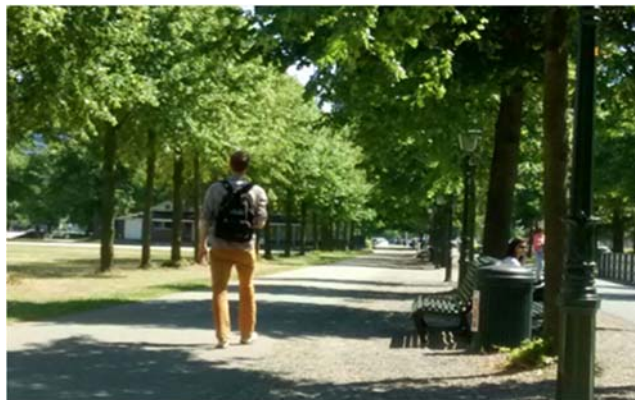
³ Photos were made after the experiment and scenes were re-enacted for illustration.

Photo 3: Helping a norm violator in the time delay treatment



Note: Notice the empty plastic bottle and the glove in the scene.

Photo 4: Characteristics of location and position of researcher



Note: The location and the position of the researcher permits private and anonymous decisions.

Photo 5: Characteristics of participants



Note: A subject must be alone and in no visible hurry nor visibly distracted. There is also no other subject coming in the opposite direction of the sidewalk

Photo 6: Location, front view



Photo 7: Location, left view



Photo 8: Location, right view



Photo 9: Location, back view



Photo 10: Example of helping behavior: participant bends down, picks up the glove and gives it back to the actor



Photo 11: Example of helping behavior: voice alert and pointing



Photo 12: Participant defects



Photo 13: Post-experimental survey



Photo 14: Close-up of the actor surveying the participant after the experiment



9.3 POST-EXPERIMENT SURVEY

After each participant crossed Point A, the actor/actress followed the participant and asked:
“Excuse me, I am a researcher of Erasmus University and I just (littered and) dropped my glove as an experiment. Could I ask you a few quick questions? We can walk together if you want.”

- C1. Did you see the (littering) and the drop of the glove ? 0. Yes _____. 1. No _____.
- C2. How willing are you to take risks in general? From 0 to 10 where max. is 10: _____.
- C3. What is your height in cm? _____ cm.
- C4. What is your age? _____ years.
- C5. How long have you lived in The Netherlands? _____ (in years / months).
- C6. How difficult was to make the decision of what to do? 0. Not at all _____. 1. Just a little _____.
- C7. Comments _____.

Note: The survey was administered in English, since The Hague is an international city and the large majority of people speaks English. In three cases respondents did not speak sufficient English and did not respond to our questions.