

**Document de travail du LEM / Discussion paper LEM  
2019-10**

# Does the presence of a physically disabled person in the group increase cooperation? An experimental test of the empathy altruism hypothesis

## Cooperation between individuals with and without disabilities

**Arnaud TOGNETTI**

Department of Clinical Neuroscience, Karolinska Institutet, Stockholm – Sweden / Institute for Advanced Study in Toulouse

**David DOAT**

Catholic University of Lille – ETHICS (EA7446)

**Dimitri DUBOIS**

CEE-M, University of Montpellier, CNRS, INRA, Montpellier SupAgro

**Rustam ROMANIUC**

LEM UMR 9221 / Catholic University of Lille – ETHICS (EA7446) / [rustam.romaniuc@univ-catholille.fr](mailto:rustam.romaniuc@univ-catholille.fr)

<http://lem.cnrs.fr/>

Les documents de travail du LEM ont pour but d'assurer une diffusion rapide et informelle des résultats des chercheurs du LEM. Leur contenu, y compris les opinions exprimées, n'engagent que les auteurs. En aucune manière le LEM ni les institutions qui le composent ne sont responsables du contenu des documents de travail du LEM. Les lecteurs intéressés sont invités à contacter directement les auteurs avec leurs critiques et leurs suggestions.

Tous les droits sont réservés. Aucune reproduction, publication ou impression sous le format d'une autre publication, impression ou en version électronique, en entier ou en partie, n'est permise sans l'autorisation écrite préalable des auteurs.

Pour toutes questions sur les droits d'auteur et les droits de copie, veuillez contacter directement les auteurs.

The goal of the LEM Discussion Paper series is to promote a quick and informal dissemination of research in progress of LEM members. Their content, including any opinions expressed, remains the sole responsibility of the authors. Neither LEM nor its partner institutions can be held responsible for the content of these LEM Discussion Papers. Interested readers are requested to contact directly the authors with criticisms and suggestions.

All rights reserved. Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorization of the authors.

For all questions related to author rights and copyrights, please contact directly the authors.

# Does the presence of a physically disabled person in the group increase cooperation? An experimental test of the empathy-altruism hypothesis

*Cooperation between individuals with and without disabilities*

Arnaud Tognetti<sup>a,b</sup>, David Doat<sup>c</sup>, Dimitri Dubois<sup>d</sup>, Rustam Romaniuc<sup>c,e</sup>

<sup>a</sup> Department of Clinical Neuroscience, Karolinska Institutet, Stockholm – Sweden

<sup>b</sup> Institute for Advanced Study in Toulouse, 21 allée de Brienne, 31015 Toulouse – France

<sup>c</sup> Catholic University of Lille – ETHICS (EA7446)

<sup>d</sup> CEE-M, University of Montpellier, CNRS, INRA, Montpellier SupAgro, Montpellier - France

<sup>e</sup> LEM (UMR 9221)

\* Corresponding author: [arnaud.tognetti@gmail.com](mailto:arnaud.tognetti@gmail.com), Karolinska Institutet, Department of Clinical Neuroscience, Division of Psychology, Nobels väg 9, 171 65 Solna

## Abstract

The empathy-altruism hypothesis postulates that the awareness of others' need, pain, or distress increases empathetic feelings, which in turn triggers cooperative behaviour. Although some evidence supports this hypothesis, previous studies were prone to the 'experimenter demand effects' raising concerns about the interpretation of the results. To avoid this issue, we designed a laboratory experiment where we examined whether the presence of individuals with a genuine physical disability would increase group cooperation in a public goods game. By manipulating the group composition during a social dilemma, we created a more ecologically valid environment closer to real-life interactions. Our results showed that the presence of physically disabled individuals did not affect group cooperation. Specifically, their presence did not affect the contributions of their physically abled partners. The lack of a surge in cooperative behaviour questions the interpretation of previous studies and suggests that they may be explained by an experimenter demand effect. Alternatively, our results may also suggest that in the context of a social dilemma with real stakes, people with physical disabilities are not perceived as being in need or do not induce enough empathy to overweight the cost of cooperation and trigger cooperative behaviours.

## Keywords

Cooperation, empathy-altruism hypothesis, public goods game, physically disabled individuals

## Acknowledgements

We are very grateful to all the participants in our study and to Cécile Donzé for helping us at the recruitment stage. We thank the CNRS for its financial support through its competitive funding programs for interdisciplinary research (project S2C3 "Social and Cognitives Sciences of the Collective Behaviors").

## Introduction

There is overwhelming evidence from laboratory and field experiments that genetically unrelated individuals cooperate despite individual costs and a lack of perceived benefits (Rand and Nowak, 2013). While many scholars argue that the evolution of cooperation in social dilemmas is determined by structural variables, including group size (Barcelo & Capraro, 2015), the marginal benefit received (Nosenzo et al., 2015), monitoring techniques (Ostrom et al., 1992), and the information available to participants (McAuliffe et al., 2019), others claim that individual psychological factors are more pertinent (Kirman & Teschl, 2010).

One potential psychological factor could be empathy, which is generally defined as the capacity to put oneself in someone else's shoes and thus be concerned for another person's position (Hoffman, 1975). The idea that empathy is a major determinant of prosocial behaviour is widely accepted among psychologists (Eisenberg & Miller, 1987; Batson & Moran, 1999; Batson & Ahmad, 2001; Batson, Lishner & Stocks, 2015) and it has furthermore been revived in economics with the development of behavioural economics (Bowles & Gintis, 2003; Camerer, 2003). For example, the empathy-altruism hypothesis postulates that empathic concern produces altruistic motivation (Batson, 1987).<sup>1</sup>

Although widely accepted, the existing empirical evidence on the effect of empathy-induction on prosocial behaviour is however sparse. Batson and Moran (1999) were one of the first to investigate empathy-induced altruism as a motive for cooperation in the context of a social dilemma. They placed 60 undergraduate women in a fictitious, one-trial prisoner's dilemma. To induce empathy, some participants read a note revealing a need for cheering up from their partners. They found that the participants in the empathy-induction condition behaved more cooperatively than the participants in the control condition.<sup>2</sup> More recently, Klimecki et al. (2016) conducted a within-subject experiment where they compared subjects' decisions in the standard one-shot dictator game with a novel version, the so-called empathic dictator game. In the latter version of the game, empathy was elicited through the presentation of videos depicting a suffering person in need: a helpless child in an orphanage. All subjects participated first in the standard dictator game, followed by the empathic version. In line with the empathy-altruism hypothesis, the authors observed significantly higher donations to unfamiliar individuals in the empathic version of the game compared to the standard one. Notwithstanding Klimecki et al.'s contribution to the literature, as well as the earlier works on the topic summarized in Batson, Lishner and Stocks (2015), the use of videos (pictures or messages) for empathy induction is a design choice that is prone to what has been called "experimenter demand effects" (EDE) (Zizzo, 2010). EDE refer to changes in behaviour by the subjects due to cues about what constitutes

1 As Batson, Lishner and Stocks (2015) note, empathy may "include feelings of sympathy, compassion, softheartedness, tenderness, sorrow, sadness, upset, distress, concern, and grief" and is other-oriented.

2 In an earlier study, Batson et al. (1995) investigated whether feeling empathy for another member of the collective in a social dilemma would create an altruistic desire to allocate resources to that person as an individual, reducing the collective good. They used the same procedure as in Batson and Moran (1999) for empathy induction.

appropriate behaviour. EDE can be social or purely cognitive and are a potential problem as they may be associated with the experimental objectives. Studies using videos or photographs of people suffering in a within-subject experiment are thus problematic since subjects can easily infer the experimental objectives and behave according to what they think the experimenters expect from them which in turn may bias the results. These effects are often connected to Milgram's (1974) experiment and to the Hawthorne factory experiments where higher productivity seemed to occur when workers were the object of a study (Mayo, 1933).

While we acknowledge that EDE cannot be entirely ruled out when studying the empathy-altruism hypothesis, we proposed an experimental design aiming minimizing the relevance or plausibility of EDE based criticism. We designed a controlled laboratory experiment where we examined whether the presence of individuals with a genuine and conspicuous physical disability would increase group cooperation in the context of a public goods game with real stakes. By manipulating the group composition during a social dilemma game, we created a more ecologically valid environment closer to real-life interactions between persons with and without visible physical disabilities reducing the likelihood for subjects to infer the experimental objectives. Further, our experimental settings enable us to test whether the presence of physically disabled individuals influences cooperativeness of physically abled individuals both after a brief exposure (mimicrying previous studies using only video or pictures of person in need) and during real cooperative interactions. The empathy-altruism hypothesis would predict that the presence of physically disabled individuals would increase cooperation by inducing higher contributions from the physically abled individuals and that this effect would be stronger during real cooperative interactions than after a brief exposure. In addition, because there is evidence within the empathy-altruism literature, that women express more empathic feelings than men (Rueckert & Naybar, 2008), we can also predict that the presence of a physically disabled person would affect more women's than men's cooperative behaviour. However, because EDE may have inflated the results of previous studies, we predict that the presence of a physically disabled person will have a less strong effect on cooperation (if any) in our setting compared to the effect of stimuli implemented in previous studies (e.g., pictures, videos, direct messages).

## **Materials & Methods**

### **Experimental method**

To study cooperation, we used what has become the benchmark for experimental research on social dilemmas: the public goods game (Ledyard, 1995). The public goods game represents a stylized model of a community in which each person's wellbeing depends on own and other persons' contributions. Individually, each member is best off if s/he contributes nothing and relies on others' efforts to create social benefits by behaving cooperatively.

Participants in our game were assigned to groups of four and endowed with 20 tokens. They had to choose how to allocate this endowment between a public account and a private account. Each token left on the private account yielded to the subject a benefit of 1 Experimental Currency Unit (ECU). In addition to the ECU kept on the private account, each group member received a fixed benefit of 0.5 ECU from each token invested in the public account by any group member. From these parameters it follows that the utilitarian optimum and the efficient symmetric outcome is for all group members to contribute their entire endowments to the public account. However, under these specifications, it nonetheless remains in each individual's self-interest to contribute zero.

At the outset of each session, participants were informed that the central server would allocate them randomly to groups of four people. We employed partner matching to mimic real-world situations in which the same group members interact repeatedly (e.g., in the workplace), so group assignments remained the same for the entire session. Each session consisted of ten periods. The number of periods was common knowledge, as was the fact that participants' decisions and identity were anonymous during the game. Each period was divided in two steps. First subjects had to decide on the allocation of the 20 tokens between their private and the public account. Second, subjects were informed about the total contribution of their group and their payoff for the current period. At the end of the ten round public goods game, participants answered a short questionnaire (age, gender, level of study, current occupation (student or not), study discipline).

In two sessions, we recruited participants from the student population at the Catholic University of Lille, France, to play ten periods of a standard public goods game in groups of four individuals, which corresponds to our baseline. Students were invited via the ORSEE software (Greiner, 2015), from a pool of more than 3,000 volunteers, to participate in an experiment at the Anthro-Lab (Catholic University of Lille). Forty subjects (57.50% males) participated in the baseline sessions. They were all physically valid and 23.58 (SD = 3.66) years old on average. None of them had previously participated in a public goods game or a game with similar parameters.

In other eight sessions, in addition to the students' population (N = 112), we recruited 16 participants with visible physical disabilities. Most of the latter who participated in our experiment suffered from multiple sclerosis. They were all in wheelchairs, which made their disability conspicuous to others. The participants with physical disabilities were recruited from local associations and at the annual meeting of the national association of persons with multiple sclerosis, which took place in Lille, France on May 2017. At the recruitment stage, they received similar information to the student population. It is worth noting that the 16 participants (50% males) with physical disability were older (M = 53.44, SD = 10.37 years) than the other participants in the same sessions (M = 22.54, SD = 6.26 years; 51% males).

In each of these eight sessions, we had two participants with visible physical disability and fourteen participants with no (visible) physical disability, which made 4 groups of 4 in each session. The

sixteen participants were randomly seated at one of the terminals (terminals were separated by lateral partitions to ensure complete anonymity), as in the baseline. They received identical instructions to the baseline. The only difference with the baseline sessions is that at the group formation stage the computer randomly allocated the two participants with visible physical disability to two different groups where the other 3 participants had no visible physical disability and this was common knowledge. That is, for each participant we displayed the following information on the participant's computer screen: "There are two groups where there is one person with physical disability. In your group, there is one person with physical disability (or alternatively, in your group there is no one with physical disability)." Hence, in each of the eight sessions, half of the groups were what we will further refer to as mixed groups (i.e. groups composed of one participant with visible physical disability and three participants with no visible physical disability), and the other half were non-mixed groups. The comparison between contribution decisions in the baseline and in the non-mixed groups will allow us to identify the effect on cooperation from a brief exposure to persons with visible physical disability before the start of the public goods game. The comparison between mixed groups and non-mixed groups will tell us whether cooperation is affected by the interaction with a physically disabled person in the group.

Overall, participants in the experiment earned an average of 14.64 € (SD = 2.57€) for one hour of participation, including initial instructions and payment of participants.

## **Statistical analyses**

We first examined whether the level of contributions to the public good differed between physically disabled and abled participants. To do this, we used a non-parametric Wilcoxon rank-sum test to compare the average contributions to the public account across the 10 rounds of the game between the two participant categories within the mixed groups. In addition, we investigated whether the dynamics of contributions over time differed between the two categories. We thus used a linear mixed model with the individual's contribution to the public account in a given round as the response variable. Our explanatory variable was the participant's category (physically disabled versus abled) and we included the variable 'round' and its interaction with the participant's category. Finally, 'participant ID' nested within 'group ID' were included as random factors to account for the individual's repeated measures and their non-independence within groups.

Second, we investigated whether the game conditions influenced cooperation in a group. We used a non-parametric Kruskal-Wallis test by ranks to examine whether average contributions across the 10 rounds differed between the three types of groups (baseline groups, non-mixed groups and mixed groups). We also used a linear mixed model to examine the dynamics of the groups' contributions over time. The response variable was the group's contribution to the public account in a given round.

Our key explanatory variable was the group's type and we include the variable 'round' and its interaction with the group's type. We also included 'group ID' as a random factor to account for the repeated measures across rounds.

Finally, we investigated whether and how game conditions influenced individual contributions to the public account of physically abled participants. Analyses were performed separately for each gender. We used a non-parametric Kruskal-Wallis test by ranks to examine whether the individual average contributions differed between the three types of groups. Then, we used linear mixed models with the individual's contribution to the public account in a given round as the response variable to investigate whether group's type influences the pattern of individual contributions across rounds. The key explanatory variable was the group's type, as in the previous model, and we also included the variable 'round' and its interaction with the group's type. The interaction was removed from the models when it did not significantly influence the response variable. We also included participant's age as a control variable. Finally, 'participant ID' nested within 'group ID' were included as random factors to account for the individual's repeated measures and their non-independence within groups.

All analyses were conducted using R 3.4.2. Linear mixed models were performed using the lmer function of the lme4 package.

## **Results**

### **Contributions of physically disabled and non-physically disabled participants**

Within the mixed groups, the average contribution to the public account across the 10 rounds significantly differed between the two participant categories (Wilcoxon rank-sum test:  $U(N = 64) = 27.0$ ,  $p = .03$ ): physically disabled participants contributed significantly more ( $M = 12.20$ ,  $SD = 5.78$  tokens,  $N = 16$ ) than physically abled participants ( $M = 9.57$ ,  $SD = 7.12$  tokens,  $N = 48$ ). This was supported by the linear mixed model (Table S1) finding a significant effect of the interaction between participants' categories and rounds ( $X^2(1, N = 640) = 26.81$ ,  $p < .0001$ ) suggesting that contributions of non-physically disabled participants decreased over time ( $\beta = -0.61$ ,  $SE = 0.07$ , 95% CI = [-0.74;-0.48]), while this is not the case for the physically disabled participants ( $\beta = 0.09$ ,  $SE = 0.11$ , 95% CI = [-0.13;-0.32]) (Fig. 1).

### **The effect of physically disabled participants on group contributions**

The average contribution to the public account was not significantly different (Kruskal-Wallis test by ranks:  $H(2, N = 42) = 1.54$ ,  $p = .46$ ) between the baseline groups ( $M = 36.37$ ,  $SD = 17.74$  tokens,  $N = 10$ ), the non-mixed groups ( $40.90 \pm 18.09$  tokens,  $N = 16$ ), and the mixed groups ( $M = 33.75$ ,  $SD = 18.90$  tokens,  $N = 16$ ). Although the linear mixed model showed a marginally significant effect

of the interaction between group's type and rounds ( $X^2(2, N = 420) = 6.04, p = .05$ ; Table S2), contributions decreased over time for all three types of groups (baseline:  $\beta = -1.86, SE = 0.32, 95\% CI = [-2.50; -1.22]$ ; non-mixed groups:  $\beta = -2.58, SE = 0.26, 95\% CI = [-3.09; -2.08]$ ; mixed groups:  $\beta = -1.74, SE = 0.26, 95\% CI = [-2.24; -1.23]$ ) (Fig. 2). Post-hoc analyses comparing contributions between the three types of groups two-by-two showed that contributions in mixed groups decreased significantly less sharply than in non-mixed groups ( $\beta = 0.84, SE = 0.36, t = 2.32, p = .05$ ). No other significant difference was found (all  $p > .18$ ).

### **The effect of physically disabled participants on individual contributions across the three conditions**

Men's (baseline groups ( $M \pm SD$ ):  $11.54 \pm 7.05$  tokens; non-mixed groups:  $10.09 \pm 7.20$  tokens, mixed groups:  $10.48 \pm 7.35$  tokens) and women's (baseline groups:  $7.29 \pm 6.21$  tokens; non-mixed groups:  $6.98 \pm 6.38$  tokens, mixed groups:  $9.95 \pm 6.38$  tokens) average contributions to the public account did not differ between the three conditions (Kruskal-Wallis test by ranks, men:  $H(2, N = 72) = 0.23, p = .89$ ; women:  $X^2(2, N = 80) = 2.46, p = .29$ ). The linear mixed models were performed without the interaction between group's type and rounds as this interaction did not significantly influence men's ( $X^2(2, N = 720) = 3.39, p = .18$ ) and women's ( $X^2(2, N = 800) = 2.74, p = .25$ ) contributions to the public account. The models (Table S3) did not find any evidence that the presence of physically disabled participants influenced men's ( $X^2(2, N = 720) = 0.83, df = 2, p = .66$ ) and women's contributions ( $X^2(2, N = 800) = 1.79, p = .41$ ) to the public account (Fig.3). Both men's and women's contributions significantly decreased over rounds (men:  $X^2(1, N = 720) = 72.17, p < .0001$ ; women:  $X^2(1, N = 800) = 130.31, p < .0001$ ). Finally, age significantly influences women's ( $X^2(1, N = 800) = 20.14, p < .0001$ ) but not men's contributions ( $X^2(1, N = 720) = 0.01, p = .96$ ).

### **Discussion**

The empathic-altruism hypothesis suggests that empathic concern produces altruistic motivation (Batson, 1987). Although there is a large range of evidence supporting the empathic-altruism hypothesis (for a review see: Batson, Lishner & Stocks, 2015), previous empirical studies conducted so far may suffer from the experimenter demand effects (EDE) potentially inflating the results obtained. By minimizing the relevance or plausibility of EDE, we experimentally tested whether the exposure to physically disabled persons would increase one's cooperative behaviours in the context of a public goods game with real stakes. Contrary to the empathic-altruism hypothesis, our results find no significant effect from the exposure to a physically disabled person on other's cooperative behaviours. Our findings question the interpretation of the results from previous studies and highlight the need to control for the experimenter demand effects in future investigations.

Our results indicate that both a brief exposure to or a real interaction with physically disabled individuals do not significantly affect the contributions to the public good of physically abled participants. In addition, the presence of physically disabled individuals did not influence the dynamics of groups' contributions: group's contribution decreased over time in the three experimental conditions, as it is classically observed in standard public goods game (Chaudhuri, 2011). Although the results did find that dynamics of groups' contributions over time differ between mixed- and non-mixed group's, this effect is merely due to physically disabled participants' behaviours. Indeed, physically disabled participants contributed significantly more to the public good than physically abled participants and their contributions did not decrease over time (this higher cooperativeness from physically-disabled participants is however difficult to interpret since the former participants are older, on average, than the physically abled ones).

Hence, our results did not find any evidence supporting the empathic-altruism hypothesis. This is surprising considering the extant literature showing that in both humans and animals the willingness to incur a cost for others to receive a benefit is higher when the others are in need, pain, or distress (Batson et al., 2015, de Waal, 2007; Klimecki et al., 2016). Previous research has also shown that compassion, which denotes a feeling of care for a suffering other accompanied by the desire to help, motivates cooperative behaviour (Saslow et al., 2013). The lack of a surge in the desire to behave more prosocially in the presence of a physically disabled person may indicate that the latter do not invoke enough feelings of compassion or empathy in economically constrained environments to overweight the cost of cooperation and increase prosocial behaviour. This observation, in turn, can be interpreted as evidence that in such environments physically disabled individuals are not considered to be more in need, pain or distress than individuals without any visible physical disability.

A different, although complementary, explanation for our results has to do with the method that we used in this experiment to study behavioural reactions to the presence of physically disabled individuals. Contrary to studies that show a positive correlation between one's exposure to a person in need, pain or distress and his/her prosocial behaviour, we did not use photographs, videos or messages to induce empathy in our participants. Indeed, such a method suffers from potential bias such as the experimenter demand effects (Zizzo, 2010). Instead, we implemented a more realistic environment with groups composed of participants with and without a conspicuous physical disability interacting in the public goods game under identical conditions, similarly to real-life interactions between these two categories. Hence, the fact we did not observe any effect of their presence questions the interpretation of previous studies for which the EDE may have biased participant's behaviours to the experimental conditions. It paves the way for further studies investigating the empathic-altruism hypothesis controlling for the EDE.

It is worth-noting that while the use of physically disabled persons renders the environment within which the participants interact more realistic since it may seem natural that these people came to the

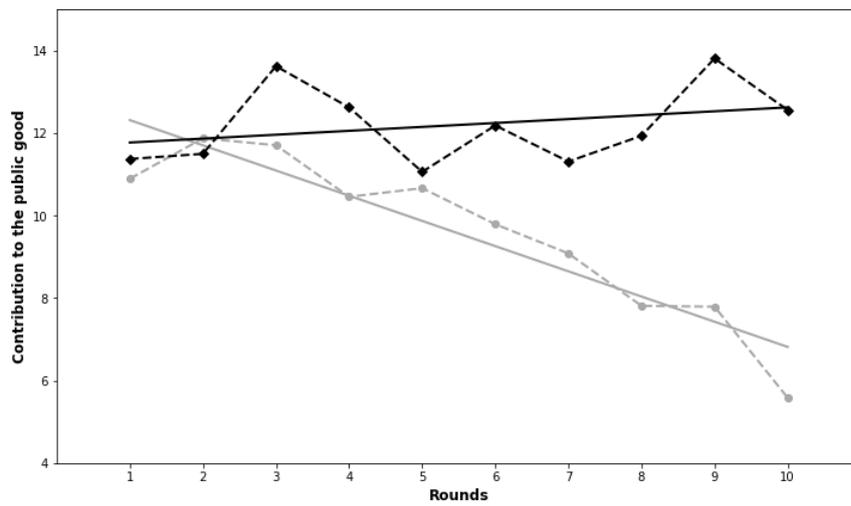
lab to participate in an experiment (which may happen on any regular day in the lab), this may also represent a limit of our study because there is a disconnection between physical disability and the need for money that can be earned if the group members behave cooperatively. However, our results are unlikely to be explained only by this disconnection as some previous studies supporting the empathic-altruism hypothesis also used economic games. In addition, there is evidence that an individual's cooperativeness during economic games is related to his/her cooperativeness in other type of settings including in daily-life contexts (Peysakhovich, Nowak & Rand, 2014; Fehr & Leibbrandt, 2011). Therefore, we expect that empathic concerns for physically disabled individuals would also trigger cooperative behaviours during economic games.

## References

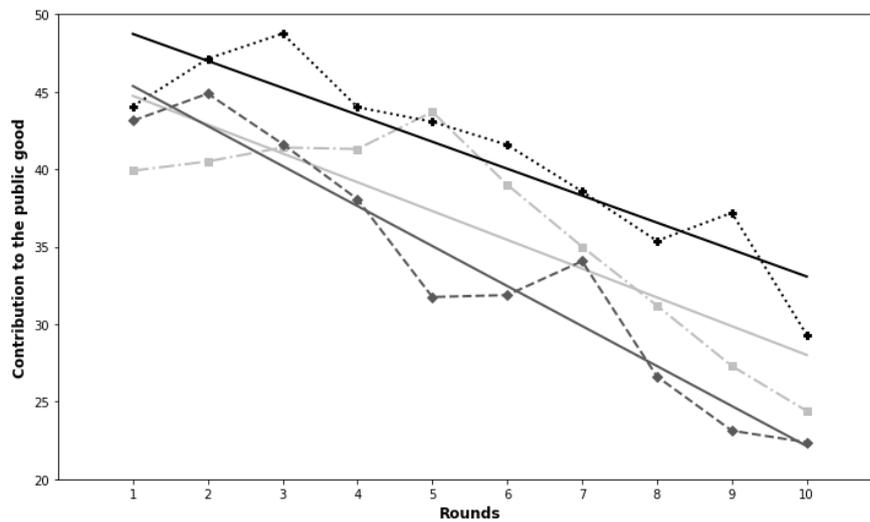
- Batson, C. D., Batson, J. G., Todd, R. M., Brummett, B. H., Shaw, L. L., Aldeguer, C. M. R. (1995). Empathy and the collective good: Caring for one of the others in a social dilemma. *Journal of Personality and Social Psychology*, 68, 619–631.
- Batson, C. D., & Moran, T. (1999). Empathy-induced altruism in a Prisoner's Dilemma. *European Journal of Social Psychology*, 29, 909–924.
- Batson, C. D., & Ahmad, N. (2009). Using empathy to improve intergroup attitudes and relations. *Social Issues and Policy Review*, 3, 141–177.
- Batson, C.D., Lishner, D.A., Stocks, E.L. (2015). The empathy-altruism hypothesis. In: Schroeder, D.A., Graziano, W.G, *The Oxford Handbook of Prosocial Behavior*. Oxford.
- Camerer, C. (2003). *Behavioral game theory: Experiments in Strategic Interaction*. Princeton University Press.
- Capraro, V., Barcelo, H. (2015). Group size effect on cooperation in one-shot social dilemmas II: Curvilinear effect. *PLoS ONE*, 10(7), doi:10.1371/journal.pone.0131419.
- Chaudhuri, A. (2011). Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature. *Experimental Economics* 14(1),47–83.
- Christov-Moore, L., Simpson, E.A., Coudé, G., Grigaityte, K., Iacoboni, M., Ferrari, P.F. (2014). Empathy: Gender effects in brain and behavior. *Neuroscience and Biobehavioral Reviews*, 46, 604-627.
- De Waal, F.B. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology*, 59, 279-300.
- Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin*, 101(1), 91-119.
- Fehr, E., Leibbrandt. (2011). A Field Study on Cooperativeness and Impatience in the Tragedy of the Commons. *Journal of Public Economics* 95, 1144–55.
- Gintis, H., Bowles, S. (2003). *A cooperative species: Human reciprocity and its evolution*. Princeton University Press.
- Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1: 114-125.
- Kirman, A., Teschl, M. (2010). Selfish or selfless? The role of empathy in economics. *Philosophical Transactions of the Royal Society: B*, 365, 303-317.
- Klimecki, O.M., Mayer, S.V., Jusyte, A., Scheeff, J., Schoenberger, M. (2016). Empathy promotes altruistic behavior in economic interactions. *Scientific Reports*, 31961.

- Ledyard, O. (1995). Public goods: some experimental results. In: Kagel, J. and Roth, A. (eds.), *Handbook of experimental economics*. Princeton: Princeton University Press (Chap. 2).
- Mayo, E. (1933). *The human problems of an industrial civilization*. New York: MacMillan.
- McAuliffe, W., McCullough, M., Burton-Chellew, M. (2019). Learning whether to cooperate in novel social interactions. PsyArXiv.
- Milgram, S. (1974). *Obedience to authority: an experimental view*. New York: Harper and Row.
- Nosenzo, D., Quercia, S., Sefton, M. (2015). Cooperation in small groups: The effect of group size. *Experimental Economics*, 18(1), 4-14.
- Ostrom, E., Walker, J., Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *American Political Science Review*, 86(2), 404-417.
- Peysakhovich, A., Nowak M., Rand D. (2014). Humans Display a “Cooperative Phenotype” That Is Domain General and Temporally Stable. *Nature Communication*, 5, 4939.
- Rand, D., Nowak, M. (2013). Human cooperation. *Trends in Cognitive Science*, 17(8), 413-425.
- Saslow, L.R., Muise, A., Impett, E.A., Dubin, M. (2013). Can you see how happy we are? Facebook images and relationship satisfaction. *Social Psychological and Personality Science*, 4(4), 411-418.
- Zizzo, D. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, 13(1), 77-98.

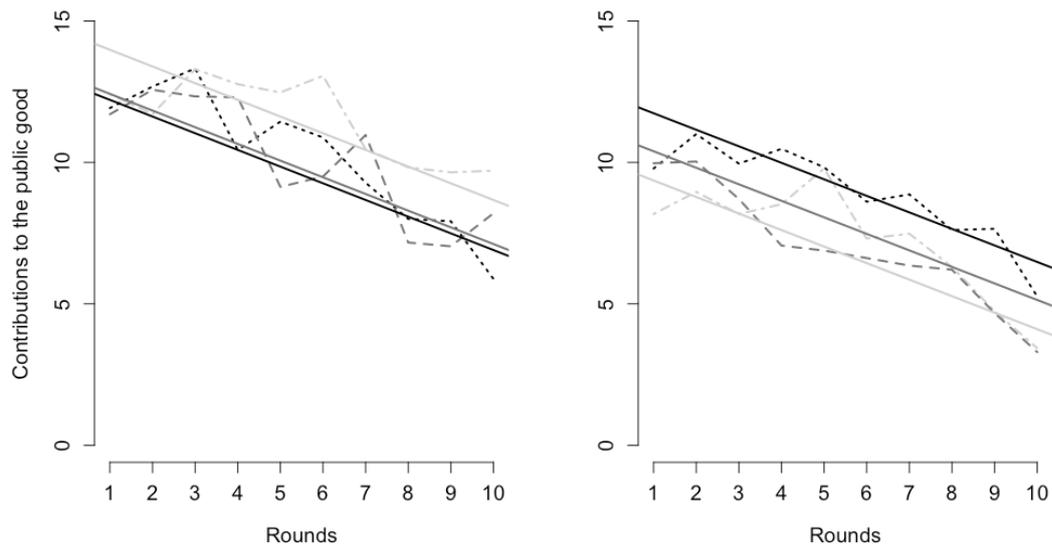
## Figures



**Fig.1.** Average contributions over the 10 rounds for the physically disabled participants (black) versus the non-physically disabled participants (grey). Raw data are shown in dotted lines. Predicted values are shown in solid lines.



**Fig. 2.** Average contributions over the 10 rounds in mixed groups (black), non-mixed groups (dark grey) and baseline groups (light grey). Raw data are shown in dotted lines. Predicted values are shown in solid lines.



**Fig. 3. Men's (left) and women's (right) average contributions to the public account over the rounds in mixed groups (black), non-mixed groups (dark grey) and baseline groups (light grey). Raw data are shown in dotted lines. Predicted values are shown in solid lines.**

## Supplemental Materials

**Table S1. Linear mixed model investigating the effect of physical disability on contributions to the public account ( $N_{\text{observation}}=640$  from 64 individuals split in 16 groups).** For each factor, the estimate, standard error of the mean (SE), degrees of freedom (df),  $\chi^2$  statistic and  $p$ -value of the likelihood ratio test of the comparison between the full model and the model without the factor, are given. For categorical variables, the estimates are for one category compared to the reference category (underlined term).  $R^{2m}$  is the variance explained by fixed factors of the model (marginal  $R^2$ ) and  $R^{2c}$  is the variance explained by both fixed and random factors (conditional  $R^2$ ). Participant's ID nested within group's ID were included as random-effect variables.

$R^{2m} = 0.07$ ; $R^{2c} = 0.63$	Estimate	SE	$\chi^2$	df	$P$ -value
(Intercept)	12.93	1.12			
Participant's category					
disabled/ <u>non-disabled</u>	-1.25	1.40			
Rounds	-0.61	0.07			
Participant's category x Round			26.81	1	<0.0001
disabled/ <u>non-disabled</u> x Round	0.71	0.14			

**Table S2. Linear mixed model investigating the effect of game condition on group contribution to the public account (N<sub>observation</sub>=420 from 42 groups).** For each factor, the estimate, standard error of the mean (SE), degrees of freedom (df),  $\chi^2$  statistic and *p*-value of the likelihood ratio test of the comparison between the full model and the model without the factor, are given. For categorical variables, the estimates are for one category compared to the reference category (underlined term). R<sup>2m</sup> is the variance explained by fixed factors of the model (marginal R<sup>2</sup>) and R<sup>2c</sup> is the variance explained by both fixed and random factors (conditional R<sup>2</sup>). Group's ID was included as a random-effect variable.

R <sup>2m</sup> = 0.13; R <sup>2c</sup> = 0.76	Estimate	SE	$\chi^2$	df	<i>P</i> -value
(Intercept)	46.59	5.19			
Group's type					
mixed/ <u>baseline</u>	3.88	6.61			
non-mixed/ <u>baseline</u>	1.36	6.61			
Rounds	-1.86	0.32			
Participant's category x Round			6.04	2	<0.05
mixed/ <u>baseline</u> x Round	0.12	0.41			
non-mixed/ <u>baseline</u> x Round	-0.72	0.41			

**Table S3. Linear mixed models investigating the effect of game condition on (A) men's (N<sub>observation</sub>=720 from 72 men split in 36 groups) and (B) women's (N<sub>observation</sub>=800 from 80 women split in 36 groups) contributions to the public account.** For each factor, the estimate, standard error of the mean (SE), degrees of freedom (df),  $\chi^2$  statistic and *p*-value of the likelihood ratio test of the comparison between the full model and the model without the factor, are given. For categorical variables, the estimates are for one category compared to the reference category (underlined term). R<sup>2m</sup> is the variance explained by fixed factors of the model (marginal R<sup>2</sup>) and R<sup>2c</sup> is the variance explained by both fixed and random factors (conditional R<sup>2</sup>). Participant's ID nested within group's ID were included as random-effect variables.

	Estimate	SE	$\chi^2$	df	<i>P</i> -value
(A) Men (R <sup>2m</sup> = 0.06; R <sup>2c</sup> = 0.48)					
(Intercept)	14.57	1.67			
Group's type			0.83	2	0.66
<u>mixed/baseline</u>	-1.77	2.05			
non-mixed/ <u>baseline</u>	-1.56	2.00			
Rounds	-0.59	0.07	72.17	1	<0.0001
Age*	<0.01	0.08	<0.01	2	0.96
(B) Women (R <sup>2</sup> = 0.16; R <sup>2c</sup> = 0.60)					
(Intercept)	9.95	1.39			
Group's type			1.79	2	0.41
<u>mixed/baseline</u>	2.38	1.81			
non-mixed/ <u>baseline</u>	1.03	1.74			
Rounds	-0.58	0.05	130.31	1	<0.0001
Age*	0.53	0.12	20.14	1	<0.0001

\* the variable 'age' was centered.