

Can you trust the good guys? Trust within and between groups with different missions

Sebastian Fehrler^{a,b,*}, Michael Kosfeld^{c,d,e,b,1}

^a Department of Political Science and Center for Comparative and International Studies (CIS), University of Zurich, Affolternstrasse 56, CH-8050 Zurich, Switzerland

^b Institute for the Study of Labor (IZA), Germany

^c Department of Management and Microeconomics, Goethe University Frankfurt, Grüneburgplatz 1, D-60323 Frankfurt am Main, Germany

^d Center for Economic Policy Research (CEPR), UK

^e Center for Economic Studies (CESifo), Germany

H I G H L I G H T S

- Subjects who identify themselves with a pro-social mission are more trustworthy than subjects who do not.
- Identification with pro-social missions comes with discrimination against out-groups.
- Out-group discrimination is much stronger than in a minimal group treatment.

A R T I C L E I N F O

JEL classification:

C72

C92

M51

Keywords:

Social identity theory

Group identity

Trust

Trustworthiness

Discrimination

Organization

A B S T R A C T

Non-governmental organizations and other non-profit organizations attract workers who strongly identify themselves with their missions. We study whether these “good guys” are more trustworthy, and how such pronounced group identities affect trust and trustworthiness within the groups and towards out-groups. We find that subjects who strongly identify themselves with a non-profit mission are more trustworthy in a minimal group setting but also harshly discriminate against out-groups when subjects are grouped by the missions they identify themselves with.

1. Introduction

Organizations have different missions. Particularly salient are missions in non-profit organizations who derive their *raison d'être* from their particular non-profit goals. In this study, we analyze the

effects of such group identities on behavior directed towards in-group members and towards out-groups.

Several studies suggest that some workers strongly care about non-profit missions (e.g., Besley and Ghatak, 2005; Delfgaauw and Dur, 2007). In a laboratory experiment, Fehrler and Kosfeld (2012) find that roughly one third of the subjects forgo a higher wage and choose a contract under which they can generate a donation to a non-governmental organization (NGO) instead. Sorting of types with different social preferences and the consequences for organizations have also been discussed in recent theoretical papers (e.g., Brekke and Nyborg, 2010; Kosfeld and von Siemens, 2011). Empirically, Brekke et al. (2011) show that sorting into groups with and without a non-profit mission leads to more cooperation within

* Correspondence to: Department of Political Science, University of Zurich, Affolternstrasse 56, CH-8050 Zurich, Switzerland. Tel.: +41 44 634 5270; fax: +41 44 634 5269.

E-mail addresses: sebastian.fehrler@uzh.ch (S. Fehrler), kosfeld@econ.uni-frankfurt.de (M. Kosfeld).

¹ Tel.: +49 69 79834823.

the first type of group in a public goods game, suggesting that sorting leads to groups with different social preferences (see also Lazear et al., 2012).

Another potentially important issue in this context, group identity, has, however, not received much attention. In-group favoritism and out-group discrimination are well known phenomena in social psychology. Even minimal group identities, induced by randomly labeling groups, can lead to intergroup discrimination. In recent years, economists have also begun to study the effects of group identities on social behavior (e.g., Charness et al., 2007; Ben-Ner et al., 2009; Chen and Li, 2009; Hargreaves Heap and Zizzo, 2009; Tsutsui and Zizzo, 2013). If group identities are strong and reflect differences in social preferences, stronger effects might be expected than in a minimal group setting.

We study trust and trustworthiness, comparing treatments with groups with minimal group identities and with groups with pro-social identities. Group identities are induced by grouping participants according to their answers to two questions in a short questionnaire that participants had to fill in before the experiment. In the minimal group treatment, subjects are grouped according to the question if they like one of the painters Paul Klee or Wassily Kandinsky, or if they like neither. In the mission treatment, subjects are grouped according to the question if they identify themselves strongly with the goals of one of the NGOs World Wildlife Fund (WWF) or Amnesty International (AI), or neither of them.

This design allows us to study whether subjects who identify themselves strongly with an NGO (the “good guys”) are more trustworthy among their peers, which would potentially make them attractive employees, and whether they discriminate against out-groups, which might cause organizational problems, e.g., if there is a mismatch between worker and leadership missions.²

2. Experimental design

At the beginning, before receiving instructions for the trust game, subjects are asked to fill in a short questionnaire on their computer screens. The questionnaire includes questions like “Do you do sports?”, “Do you play an instrument?”, and the question “Do you strongly identify yourself with the goals of one of the NGOs Amnesty International or the WWF?”. The last question is the one we use in our mission treatment. It has the following answer options: “WWF”, “Amnesty International”, and “None of the two”. One option has to be checked and multiple answers are ruled out. In the minimal group treatment, we use a different question from the same questionnaire to form groups: “Do you like one of the painters: Paul Klee or Wassily Kandinsky?”, with answer options “Klee”, “Kandinsky”, and “None of the two”. With this treatment, we relate to the classic social psychology study in this field by Tajfel et al. (1971), in which preferences about Klee and Kandinsky are used as well to form “minimal” groups. The questionnaire is designed to give the subjects the impression that they take part in a small socioeconomic survey to make it unlikely that they expect that their answers play a role in the experiment.

After reading the instructions and a short comprehension quiz, subjects play a trust game (Berg et al., 1995) in which transfer choices are limited to four options. Half of the subjects are trustors the other half trustees. All recipients receive an initial endowment of 12 points. Trustors can transfer 0, 4, 8, or 12 points to the trustee. The transfers are tripled. The trustees can then send back any integer amount of points from the points they dispose of back to the trustor.

² The related problem of worker–leadership mission mismatches for worker motivation is discussed in Besley and Ghatak (2005).

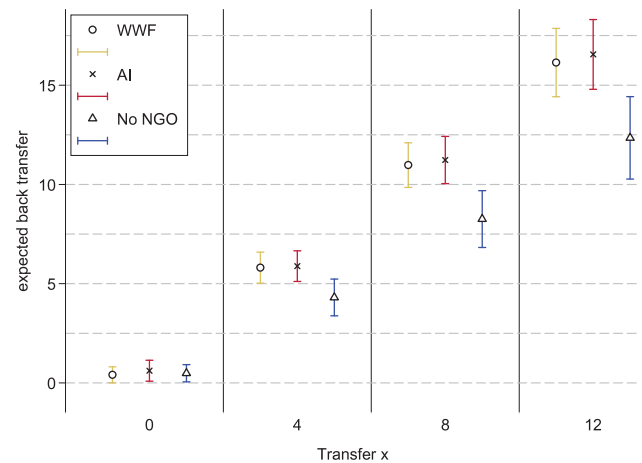


Fig. 1. Beliefs about the trustworthiness of different trustees. Note: Expected back transfers from different NGO types in the mission treatment (and their 95% confidence intervals).

Trustors and trustees can make their transfer decision in the investment game conditional on the type of the recipient, i.e., on the answer of their partner to the NGO question in the mission treatment and on the answer to the art question in the minimal group treatment. The strategy method is used. Trustors make three transfer decisions, one for each potential type of trustee. Trustees make 12 decisions, one for every possible type of trustor and received transfer.³ In addition to the transfer decisions, we ask the trustors about their beliefs regarding back transfers for all possible transfer levels and types of trustee. The answers to these questions have no influence on the pay-offs. After the transfers are made and the beliefs elicited, the experiment ends, and the subjects are paid out.

One point in the trust game is worth 0.8 CHF (at the time of the experiment, 1 CHF was worth 0.9 USD). Overall, 190 subjects (52% female) participated in the experiment in the laboratory of the Department of Economics at the University of Zurich.⁴ On average, the participants earned 14.8 CHF in addition to a show-up fee of 10 CHF, and they spent around 45 min in the laboratory.

3. Results

3.1. Trustor behavior

Fig. 1 presents the expected back transfers from different types of trustee, i.e., beliefs about their trustworthiness. We see that trustors expect lower back transfers from subjects who do not strongly identify themselves with the goals of either NGO (henceforth called No-NGO types). Regressing expected back transfers from each group on the transfers (i.e., estimating linear fits for the three groups in Fig. 1) results in statistically significantly steeper slopes for WWF and AI than for No-NGO ($p < 0.01$, F -Test, Regression (1), Table 2).⁵

Moreover, we see that the beliefs about back transfers from AI and WWF types are almost the same. Table 1 shows the transfer

³ The use of the strategy method in an investment game has been shown to lead to lower trustworthiness as compared to the “direct response” method (Casari and Cason, 2009). In the context of this experiment, this might lead to an attenuation of the effect of group identity on trustworthiness. Having subjects make transfers to the different groups of recipients appears natural when the goal is to study whether they discriminate between these groups.

⁴ The treatments were programmed with zTree (Fischbacher, 2007).

⁵ This finding holds for all types of trustor (see models (2)–(4) in Table 2).

Table 1
Transfer levels from members of each group to all different trustees.

Mission treatment				
	To WWF	To AI	To No-NGO	N
From WWF	8.3 (0.8)	7.9 (0.8)	4.6 (0.9)	28
From AI	7.1 (1.2)	8.3 (1.2)	4.6 (1.8)	14
From No-NGO	6.8 (0.7)	7.6 (0.7)	5.9 (0.8)	36
Minimal group treatment				
	To Klee	To Kandinsky	To No-Artist	N
From Klee	8.6 (1.0)	6.6 (1.1)	6.4 (1.1)	22
From Kandinsky	7.6 (0.9)	8.8 (0.8)	6.7 (1.1)	19
From No-Artist	6.3 (1.0)	6.2 (1.0)	8.3 (0.9)	26

Note: Standard errors in parentheses.

levels to the different types of trustee from the different types of trustor. The differences between the transfer levels reflect the beliefs about the back transfers. Even the No-NGO types transfer less to other No-NGO types than to AI or WWF types. For the No-NGO types, the differences of the transfer levels to the three trustee types are pairwise statistically different ($p < 0.05$, Wilcoxon rank sum test). For the other two groups, the transfer level to No-NGO types is statistically different ($p < 0.05$) to the transfer to the other two groups, which themselves are not significantly different from each other ($p > 0.1$). Transfers to No-NGO types are lower than to any other group. The NGO types receive, on average, 47% higher transfers than No-NGO types. The lower half of Table 1 shows the transfer levels in the minimal group treatment. Here, each type of trustor favors trustees with the same art preferences, but there is no single group that is less trusted than all the other groups.

3.2. Trustee behavior

In the analysis of trustee behavior, we start by looking at the trustworthiness of the different NGO types in the minimal group treatment where they cannot condition their back transfer on the NGO type of the trustor. This allows us to see whether the NGO types are more trustworthy when the group identity is unrelated to their NGO identification. In the minimal group treatment, the transfers have to be conditioned on the art preferences of the trustor. As we used the same questionnaire for both the minimal group and mission treatments, we can group the results by the answers to the NGO question. Fig. 2 presents the back transfers averaged over the three potential recipient types (“Klee”, “Kandinsky”, and “No-Artist”) for all potential transfers.

We see that people who identify themselves with one of the NGOs are more trustworthy than people who do not. Regressing back transfer on transfer gives significantly different slopes for the AI group than for the No-NGO group ($p < 0.05$, *F*-Test, Regression 5, Table 2).⁶ Pooling the AI and the WWF groups in the regression gives a significantly different slope of this combined NGO group to the No-NGO group slope ($p < 0.05$, *F*-Test, Regression 6, Table 2). The slope of the WWF group alone is not significantly different to the slopes of the No-NGO group ($p > 0.1$, Regression 5, Table 2) and the AI group ($p > 0.1$).

Next, we study what happens if subjects are grouped by their NGO identification. Fig. 3, which shows the estimated coefficients from regressing back transfers on transfers for the different groups, gives a clear picture. We see that the NGO types strongly discriminate against No-NGO types. Regressing back transfer on transfer

⁶ In this and following regressions, there are four observations from every trustor (one for each possible transfer level). In the estimation of the standard errors these are, therefore, treated as one cluster each.

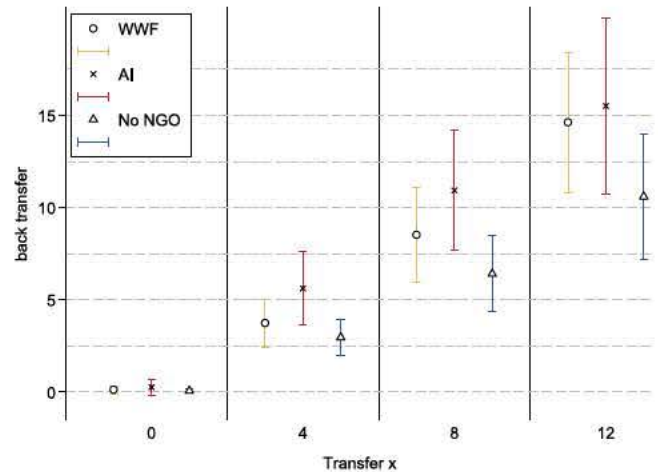


Fig. 2. Average back transfers of different NGO types in the minimal group treatment. Note: The back transfers from the different NGO groups are averaged over the three recipient (artist) groups in the minimal group treatment. The intervals displayed are 95% confidence intervals.

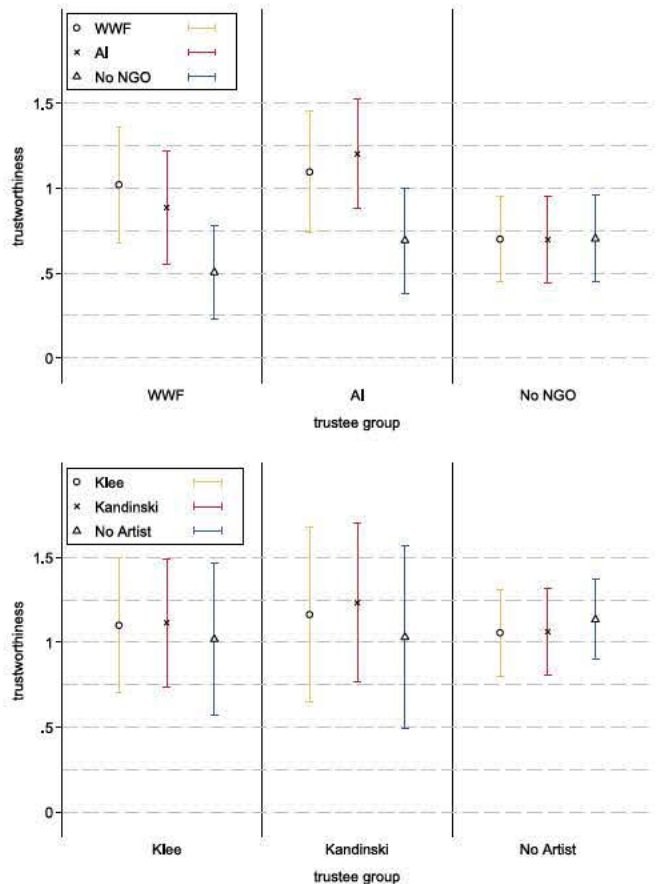


Fig. 3. Trustworthiness of different trustee groups towards different trustor groups. Note: Each panel displays the coefficients (and their 95% confidence intervals) of regressions of back transfers from a different group of trustees on transfers from a different group of trustors. Full results from the regression models are reported in Table 3.

by trustor type gives significantly different coefficients for AI and WWF types than for No-NGO types ($p < 0.05$, *F*-Test, Regressions 7 and 8, Table 3). It is also the case that WWF types favor other WWF types over AI types ($p < 0.05$), and the latter are less trustworthy to WWF types than to other AI types (but not significantly so: $p > 0.05$). These effects are small relative to the difference be-

Table 2
Regressions of expected back transfer and average back transfer on transfers from different groups.

	(1)	(2)	(3)	(4)	(5)	(6)
		WWF	AI	No-NGO		
	Dependent variable: expected back transfers in models (1)–(4) and back transfers in (5)–(6)					
WWF[0, 1]	0.48 [*] (0.23)	0.73 (0.48)	0.91 (0.57)	0.12 (0.24)	−0.49 (0.26)	
AI[0, 1]	0.60 [*] (0.28)	0.52 (0.46)	1.31 (0.82)	0.38 (0.39)	0.43 (0.29)	
No-NGO[0, 1]	0.42 (0.24)	0.28 (0.34)	0.16 (0.21)	0.63 (0.43)	−0.25 (0.20)	−0.25 (0.20)
NGO[0, 1]						−0.15 (0.21)
Transfer × WWF	1.31 ^{***} (0.073)	1.31 ^{***} (0.11)	1.19 ^{***} (0.18)	1.36 ^{***} (0.12)	1.21 ^{***} (0.15)	
Transfer × AI	1.33 ^{***} (0.075)	1.31 ^{***} (0.13)	1.28 ^{***} (0.22)	1.36 ^{***} (0.10)	1.28 ^{***} (0.17)	
Transfer × No-NGO	0.99 ^{***} (0.080)	0.94 ^{***} (0.13)	0.84 ^{***} (0.16)	1.08 ^{***} (0.13)	0.88 ^{***} (0.14)	0.88 ^{***} (0.14)
Transfer × NGO						1.23 ^{***} (0.12)
<i>N</i>	936	336	168	432	268	268
<i>R</i> ²	0.76	0.80	0.76	0.74	0.70	0.70

Note: Standard errors in parentheses. In models (1)–(2), expected back transfer is regressed on the group dummies of the trustors and their transfers. Model (1) refers to Fig. 1. In models (5) and (6), the dependent variable is the average back transfer to the three artist types. Model (5) refers to Fig. 2. In model (6), both NGO groups are pooled.
^{*} $p < 0.05$.
^{**} $p < 0.01$.
^{***} $p < 0.001$.

Table 3
Regressions of back transfers of different groups on transfers from different groups.

	(7)	(8)	(9)	(10)	(11)	(12)	
	WWF	AI	No-NGO	Klee	Kandinsky	No-Artist	
	Dependent variable in all models: back transfers						
WWF[0, 1]	1.00 (0.56)	−0.35 (0.33)	−0.12 (0.17)	Klee[0, 1]	0.58 (0.47)	0.0071 (0.15)	−0.40 (0.24)
AI[0, 1]	0.97 (0.53)	−0.050 (0.32)	−0.083 (0.19)	Kandinsky[0, 1]	−0.041 (0.34)	−0.21 (0.20)	−0.45 [*] (0.22)
No-NGO[0, 1]	0.27 (0.33)	0.017 (0.17)	−0.063 (0.21)	No-Artist[0, 1]	0.012 (0.30)	0.050 (0.14)	−0.41 (0.24)
Transfer × WWF	1.02 ^{***} (0.17)	1.09 ^{***} (0.18)	0.70 ^{***} (0.13)	Transfer × Klee	1.10 ^{***} (0.19)	1.16 ^{***} (0.24)	1.06 ^{***} (0.13)
Transfer × AI	0.88 ^{***} (0.16)	1.20 ^{***} (0.16)	0.70 ^{***} (0.13)	Transfer × Kandinsky	1.11 ^{***} (0.18)	1.23 ^{***} (0.22)	1.06 ^{***} (0.13)
Transfer × No-NGO	0.50 ^{***} (0.14)	0.69 ^{***} (0.15)	0.70 ^{***} (0.13)	Transfer × No-Artist	1.02 ^{***} (0.21)	1.03 ^{**} (0.25)	1.13 ^{***} (0.12)
<i>N</i>	344	348	524	<i>N</i>	204	168	432
<i>R</i> ²	0.37	0.48	0.37	<i>R</i> ²	0.65	0.65	0.69

Note: Standard errors in parentheses. In models (7)–(9), back transfers from the different NGO groups are regressed on the group dummies of the trustors and their transfers. Model (7) refers to the upper left panel in Fig. 3, model (8) to the upper middle panel, and model (9) to the upper right panel. In models (10)–(12), back transfers from the different artist groups are regressed on the group dummies of the trustors and their transfers. Model (10) refers to the lower left panel, model (11) to the lower middle panel, and model (12) to the lower right panel.

^{*} $p < 0.05$.
^{**} $p < 0.01$.
^{***} $p < 0.001$.

tween back transfers to the own NGO group and back transfers to the No-NGO group. They might be explained by the mere in-group effect which is present even when group formation is arbitrary as in the artist treatment. The lower half of Fig. 3 shows that there are small differences between the slopes for the different artist types. All differences except for one (back transfers to Kandinsky and to No-Artist from Kandinsky trustees) are insignificant ($p > 0.1$, F -tests, Regressions 10–12, Table 3). An explanation of the relatively small discrimination between the two NGO groups might also be that these groups could be seen as similar, and many WWF types also identify themselves with AI and vice versa (subjects could only check one group in the questionnaire). We also find that No-NGO types do not discriminate between the trustor groups (third panel in upper half of Fig. 3 and Regression 9 in Table 3).

The comparison of the behavior of the NGO types in the two treatments reveals a strong negative discrimination against out-groups and no positive discrimination of in-groups. A regression of back transfers from WWF or AI types on transfers from each group in the mission treatment and from trustors in the minimal group treatment (the three artist groups pooled) gives us four coefficients (Regressions 13 and 14, Table 4). We measure discrimination of a group in the mission treatment as the difference between trustworthiness to that group (the coefficient of the group dummy interacted with the received transfer, e.g. 1.2 for transfer × AI in Regression 14, Table 4) and trustworthiness to the artist groups in the control treatment (e.g., the coefficient 1.28 for transfer × art[avg] in Regression 14, Table 4). Discrimination is statistically not significant for WWF and AI types ($p > 0.1$, F -tests).

Table 4

Regressions of back transfers from NGO types on transfers from different groups in both treatments.

	(13) WWF	(14) AI
	Dependent variable: back transfers	
WWF[0, 1]	1.00 (0.57)	-0.35 (0.34)
AI[0, 1]	0.97 (0.53)	-0.05 (0.32)
No-NGO[0, 1]	0.27 (0.33)	0.02 (0.17)
Art[0, 1]	-0.49 (0.26)	0.43 (0.29)
Transfer × WWF	1.02*** (0.17)	1.09*** (0.18)
Transfer × AI	0.88*** (0.16)	1.20*** (0.16)
Transfer × No-NGO	0.50*** (0.14)	0.69*** (0.15)
Transfer × art[avg]	1.21*** (0.15)	1.28*** (0.17)
<i>N</i>	344	348
<i>R</i> ²	0.62	0.66

Note: Standard errors in parentheses. In models (14) and (15), back transfers from AI and WWF types, respectively, in both treatments are the dependent variable.

· $p < 0.05$.

** $p < 0.01$.

*** $p < 0.001$.

However, discrimination of the No-NGO group (-0.71 in Regression 13 and -0.59 in Regression 14) is significant ($p < 0.05$, *F*-tests). This suggests that negative discrimination of out-group subjects is the main driver of the differences in back transfers in the mission treatment.

4. Conclusion

We find that subjects who identify themselves with non-profit missions are more trustworthy if they interact with somebody with the same mission, or if they interact in a neutral setting in which they do not know the (potential) pro-social mission of their partner. Attracting such individuals might be beneficial for organizations in the non-profit sector, and may possibly explain different compensation schemes, such as fixed salaries, as compared to more performance-based remuneration schemes in for-profit organizations (e.g., Ballou and Weisbrod, 2003). However, when the group identity of the organization is salient and there is an out-group that does not share it, strong discrimination might result. A reason for the strong discrimination we observe might be that the groups differ in status (see, e.g., Tsutsui and Zizzo, 2013). The "good" NGO

groups could be seen as high-status groups and the No-NGO group as a low-status group. For an assessment of the overall welfare effects of strong identity groups as compared to a setting without any groups, future studies with a setup as in Hargreaves Heap and Zizzo (2009), who study the welfare effects of minimal groups, might be interesting.

Besley and Ghatak (2005) discuss possible detrimental effects on worker motivation if an organization hires a new principal who is not dedicated to the organization's mission. This would, for example, be the case if an NGO hires a financial expert who has gained no merits as an activist. Our results suggest that such a mission mismatch might also be detrimental for trust inside the organization. It seems that strong identification with a good cause goes hand in hand with intolerance towards out-groups.

References

- Ballou, J., Weisbrod, B., 2003. Managerial rewards and the behavior of for-profit, governmental, and nonprofit organizations: evidence from the hospital industry. *Journal of Public Economics* 87 (9–10), 1895–1920.
- Ben-Ner, A., McCall, B.P., Stephane, M., Wang, H., 2009. Identity and in-group/out-group differentiation in work and giving behaviors: experimental evidence. *Journal of Economic Behavior & Organization* 72 (1), 153–170.
- Berg, J., Dickhaut, J., McCabe, K., 1995. Trust, reciprocity, and social history. *Games and Economic Behavior* 10 (1), 122–142.
- Besley, T., Ghatak, M., 2005. Competition and incentives with motivated agents. *American Economic Review* 95, 616–636.
- Brekke, K.A., Hauge, K.E., Lind, J.T., Nyborg, K., 2011. Playing with the good guys. A public good game with endogenous group formation. *Journal of Public Economics* 95 (9–10), 1111–1118.
- Brekke, K.A., Nyborg, K., 2010. Selfish bakers, caring nurses? A model of work motivation. *Journal of Economic Behavior & Organization* 75 (3), 377–394.
- Casari, M., Cason, T.N., 2009. The strategy method lowers measured trustworthy behavior. *Economics Letters* 103 (3), 157–159.
- Charness, G., Rigotti, L., Rustichini, A., 2007. Individual behavior and group membership. *American Economic Review* 97 (4), 1340–1352.
- Chen, Y., Li, S.X., 2009. Group identity and social preferences. *American Economic Review* 99 (1), 431–457.
- Delfgaauw, J., Dur, R., 2007. Signaling and screening of workers' motivation. *Journal of Economic Behavior & Organization* 62 (4), 605–624.
- Fehrler, S., Kosfeld, M., 2012. Pro-social missions and worker motivation: an experimental study. Discussion Paper, Institute for the Study of Labor, IZA.
- Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10 (2), 171–178.
- Hargreaves Heap, S.P., Zizzo, D.J., 2009. The value of groups. *American Economic Review* 99 (1), 295–323.
- Kosfeld, M., von Siemens, F.A., 2011. Competition, cooperation, and corporate culture. *The RAND Journal of Economics* 42 (1), 23–43.
- Lazear, E.P., Malmendier, U., Weber, R.A., 2012. Sorting in experiments with application to social preferences. *American Economic Journal: Applied Economics* 4 (1), 136–163.
- Tajfel, H., Billig, M., Bundy, R., Flament, C., 1971. Social categorization and intergroup behaviour. *European Journal of Social Psychology* 1 (2), 149–178.
- Tsutsui, K., Zizzo, D., 2013. Group status, minorities and trust. *Experimental Economics* 1–30.