# Powering up with indirect reciprocity in a large-scale field experiment

Erez Yoeli[a], Moshe Hoffman[b,c], David G. Rand[c,d], and Martin A. Nowak[c,e,f,1]

[a]Federal Trade Commission, Washington, DC 20580; [b]Department of Computer Science and Engineering, University of California at San Diego, La Jolla, CA 92093; [c]Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138; [d]Department of Psychology, Yale University, New Haven, CT 06520; and Departments of [e]Mathematics and [f]Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138

Edited by John C. Avise, University of California, Irvine, CA, and approved April 3, 2013 (received for review February 14, 2013)

**A defining aspect of human cooperation is the use of sophisticated indirect reciprocity. We observe others, talk about others, and act accordingly. We help those who help others, and we cooperate expecting that others will cooperate in return. Indirect reciprocity is based on reputation, which spreads by communication. A crucial aspect of indirect reciprocity is observability: reputation effects can support cooperation as long as peoples' actions can be observed by others. In evolutionary models of indirect reciprocity, natural selection favors cooperation when observability is sufficiently high. Complimenting this theoretical work are experiments where observability promotes cooperation among small groups playing games in the laboratory. Until now, however, there has been little evidence of observability's power to promote large-scale cooperation in real world settings. Here we provide such evidence using a field study involving 2413 subjects. We collaborated with a utility company to study participation in a program designed to prevent blackouts. We show that observability triples participation in this public goods game. The effect is over four times larger than offering a $25 monetary incentive, the company's previous policy. Furthermore, as predicted by indirect reciprocity, we provide evidence that reputational concerns are driving our observability effect. In sum, we show how indirect reciprocity can be harnessed to increase cooperation in a relevant, real-world public goods game.**

evolutionary game theory | experimental economics

Cooperation occurs when we take on costs to benefit the greater good. By this definition, everyone is better off when everyone cooperates, but self-interest undermines cooperation and leads to free-riding. Promoting cooperation is a central challenge for human societies, both today and over our evolutionary history (1–10). There are five mechanisms for the evolution of cooperation (11): direct and indirect reciprocity, spatial selection, group selection, and kin selection. Each of these mechanisms is an interaction structure that can lead cooperators to outperform noncooperators, and therefore be favored by selection.

Direct and indirect reciprocity involve repeated interactions, creating future consequences for one's actions: it can pay to cooperate today to receive cooperation from others tomorrow. Spatial selection occurs when players' interactions are structured rather than occurring at random. As a result, cooperators may be more likely to interact with other cooperators and thus preferentially receive the benefits of cooperation. Spatial selection operates when cooperaters cluster in physical space, on social networks, in sets, or in phenotype space (12). Group selection (or multilevel selection) occurs when competition and reproduction happen at multiple levels: not only do players compete with others in their group, but groups compete with each other. If cooperative groups outcompete noncooperative groups, then group-level selection can favor the evolution of cooperation. Finally, kin selection may be defined as preferring to cooperate with those who are closely related. Kin recognition can allow players to cooperate with close genetic relatives and defect otherwise.

Most of the literature on the evolution of cooperation uses the Prisoner's Dilemma and related frameworks: players can pay a cost to give a greater benefit to one or more others. Thus, within the context of these games, cooperation is good for everyone. However, cooperation need not be good for everyone more generally (6, 13). There are situations in which cooperating may give a benefit to some, but impose costs on others. For example, in intergroup conflict and war, people cooperate with members of their own group in an attempt to harm members of other groups (14). Or in the context of markets, companies may collude to keep prices high, benefiting each other but harming consumers (15). The five mechanisms for the evolution of cooperation may promote both total welfare-enhancing cooperation, as well as these more pernicious forms of cooperation.

All of these mechanisms are relevant for the evolution of human cooperation, but direct reciprocity and indirect reciprocity occupy a central place: most of our key interactions are repeated and reputation is usually at stake. Direct reciprocity is based on repeated encounters between the same two individuals: my behavior toward you depends on what you have done to me. Indirect reciprocity is based on repeated encounters in a group of individuals: my behavior toward you also depends on what you have done to others (Fig. 1). We take a keen interest in who does what to whom and why, which requires sophisticated social intelligence. We talk to each other about others. As David Haig said: "For direct reciprocity you need a face, for indirect reciprocity you need a name" (4). The evolution of indirect reciprocity is linked to the evolution of human language. Supported by human language, reputation systems allow us to track the good and bad behavior of others and to use this information to incentivize cooperation. Whatever is specifically human about our mental machinery is derivative of human language, social intelligence, and thus indirect reciprocity (4, 16).

The evolution of cooperation via indirect reciprocity has been a topic of great interest in recent years. Mathematical models and computer simulations have demonstrated the power of indirect reciprocity for promoting cooperative behavior (17–40). In these models, players typically engage in a series of one-shot interactions with others selected at random from the population. In some of those interactions, players' previous decisions are observable by their partners. Observability allows players to use conditional strategies that base their actions on the partner's behavior in

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, "In the Light of Evolution VII: The Human Mental Machinery," held January 10–12, 2013, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and audio files of most presentations are available on the NAS Web site at www.nasonline.org/evolution_vii.

Author contributions: E.Y. designed research; E.Y. performed research; E.Y., M.H., D.G.R., and M.A.N. analyzed data; and E.Y., M.H., D.G.R., and M.A.N. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: martin_nowak@harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1301210110/-/DCSupplemental.

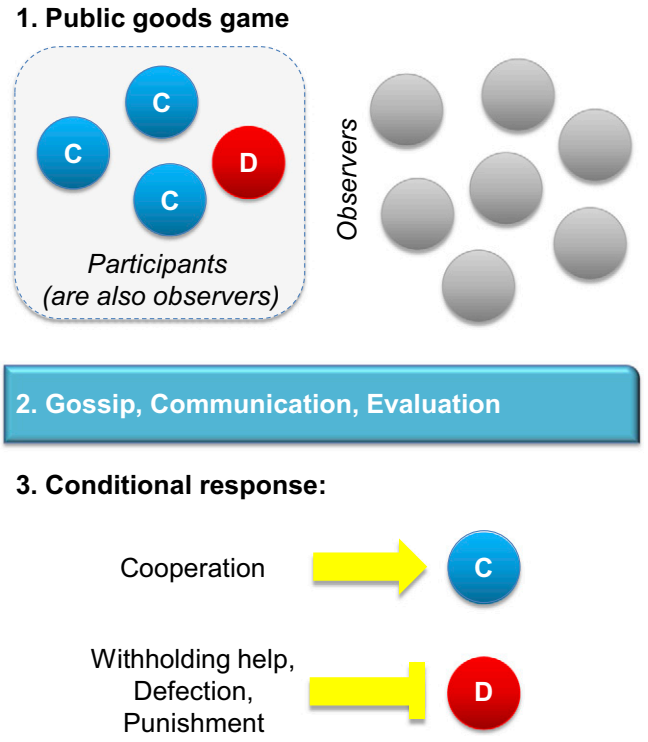



**Fig. 1.** Indirect reciprocity can support contributions to public goods, as depicted here. In indirect reciprocity, my behavior toward you depends on how you have behaved toward both myself and others. This process occurs in three stages. First, people engage in a public goods game (cooperators in blue and defectors in red). Peoples' behavior is observed, both by other players and third parties. Second, information about this behavior can spread from person to person. Based on the information received, peoples' opinions about the players are updated. Third, as a result, public goods cooperators will receive cooperation in future interactions whereas defectors will be denied cooperation, defected upon, or punished. Thus, indirect reciprocity creates an incentive to contribute to public goods and can promote the evolution of cooperation.

the past. When past actions are sufficiently observable, natural selection can favor strategies that cooperate as long as the partner has behaved well in the past. What constitutes "good" behavior worthy of receiving cooperation depends on the social norm. A simple social norm is called "image scoring" and prescribes cooperating with those who have cooperated in sufficiently many previous interactions (20). More complicated norms also take into account the behavior of the partner's previous partners. For example, under the "standing" norm, players can maintain their good reputation by defecting against those with bad reputation (24). The many models of indirect reciprocity differ in their details, yet across a wide range of assumptions, making previous decisions observable allows cooperators to selectively target their cooperation at other cooperators and withhold cooperation from defectors. Thus, free-riders are at a disadvantage, and cooperation can spread.

This body of theoretical work is supported by behavioral experiments where subjects play economic games in the laboratory. People are substantially more cooperative when their decisions are observable and when others can respond accordingly (41–60). Subjects understand that having a good reputation is valuable in these settings (49) and so are willing to pay the cost of cooperation. Observability particularly increases cooperation when the prosocial nature of the cooperative choice is made salient (55, 61). Moreover, experimental evidence indicates that indirect reciprocity is deeply entrenched in human psychology: subtle cues of observability have large effects on cooperation levels (62–65), and our initial impulse to cooperate in one-shot anonymous settings (66–69) is likely the result of adaptation in a world dominated by reputational concerns (66, 68).

These laboratory experiments are extremely valuable. They generate powerful insights into human psychology and provide clear evidence for the importance of indirect reciprocity. To do so, however, they typically use abstract economic games and involve the interaction of only a handful of subjects. Thus, the question of whether observability affects large-scale cooperation in real world settings outside of the laboratory remains largely unexplored (exceptions include refs. 70–72). The extent to which findings from theory and the laboratory generalize to natural field settings is of great importance, both for scientific understanding and for public policy (73).

Here, we address this question by running a large-scale field experiment on the effect of observability in a public goods game (PGG). We collaborated with a major electric utility company to enroll consumers in a "demand response" program. This program is designed to help prevent blackouts by reducing excessive use of air conditioning during periods of high electricity demand. The cost of electricity production can spike hundredsfold during demand peaks. However, the price consumers pay is typically constant across time. Thus, during peak periods there is a dramatic mismatch between price and actual cost, leading to excessive energy use. This mismatch reduces grid reliability, drives up energy costs, increases the risk of black outs, and harms the environment. In recent years, reducing excessive peak energy use has become a target of regulatory efforts to increase efficiency in the electricity industry.

Encouraging participation in demand response programs such as the one used in this study is the primary policy tool available for reducing peak energy use (74). Demand response programs are voluntary programs in which people allow their utility to remotely restrict their energy consumption during peak hours. To do so, the utility usually installs a remote switch in-line with the circuitry of an appliance such as a hot water heater or air conditioner. Estimates suggest that these voluntary programs could reduce the need to invest in additional generation capacity by at least 38% over the next two decades, generating cost savings of at least $129 billion (75). Voluntary energy efficiency and demand response programs have been widely available for years, but participation is frustratingly low (76). Demand response programs exemplify the public goods dilemma: participation helps reduce on-peak demand, benefitting all energy grid users, but energy consumers find participating inconvenient. Participation is socially optimal because the inconvenience is minimal for most individuals relative to the societal costs of a black out.

To explore the effect of observability on this real-world public goods problem, we solicited residents of 15 homeowners associations (HOAs) to participate in a demand response program. Residents who volunteered for this program allowed the utility to install a device that remotely curbs their central air conditioners when necessary: on days with unusually high demand or in the case of an unexpected plant or transmission failure. Residents who volunteered, therefore, contributed to a public good by improving the stability of the electrical grid in all of California, at the cost of some personal inconvenience. We solicited volunteers by delivering mailers to residents and asking them to participate. Sign-up sheets were posted in a communal area near their home, usually by a shared mailbox kiosk. In our primary manipulation, we varied whether residents' neighbors could tell who had signed up for the program. We did so by varying whether the publicly posted sheets required residents to print their name and unit number (observable treatment) or only a code that does not reveal their identity (anonymous treatment).

## Results

We found that residents in the observable treatment are nearly three times as likely to participate in the demand response program as residents in the anonymous treatment (fraction of residents participating: anonymous = 0.030, observable = 0.088, $P < 0.01$, $n = 1408$; Fig. 2). All statistics presented are from probit regressions including various controls, with SEs clustered at the HOA level; for details and regression tables, see *Supporting Information*.

The effect of the observable treatment was over seven times that of offering a $25 incentive (the estimated effect of the incentive is 0.009; a Wald test rejects that the coefficients on observability and the $25 incentive are identical, $P = 0.024$). This incentive was what the utility had used before the experiment, and they had previously argued the incentive would be far more effective than observability. In fact, this incentive appears to have been too small to be effective, and such small financial incentives are known to sometimes backfire (77). For the sake of comparison, we followed convention and estimated how large the financial incentive would have to be to achieve the same results if its effect is linear (78). We found that the utility would have had to offer an incentive of $174 to increase participation as much as our observable treatment.

We now explore the mechanism through which observability functions to increase participation. Indirect reciprocity theory is based on reputational concerns: when groups of people interact repeatedly and actions are observable, it becomes advantageous to be seen contributing to public goods. Based on this account, we predict that observability will have a greater effect among populations where ongoing relationships and reputations are expected to play a larger role. We evaluate this prediction in two ways.

First, we test whether the effect of the observable treatment was greater in apartment buildings compared with row houses and individual homes. In apartment buildings, residents are more likely to interact with their neighbors in public spaces, and sign-up sheets were typically posted in especially conspicuous locations. Thus, indirect reciprocity theory predicts that observability will have a larger effect in apartment buildings. As shown in Fig. 3*A*, the results confirm this prediction: observability increased participation among those living in apartment buildings (fraction of residents participating: anonymous = 0.048, observable = 0.114, $P < 0.01$, $n = 582$) whereas it had little effect on the inhabitants of row houses or individual homes (fraction of residents participating: anonymous = 0.024, observable = 0.038, not significant, $n = 826$; yielding an estimated interaction of 0.052, $P = 0.04$).

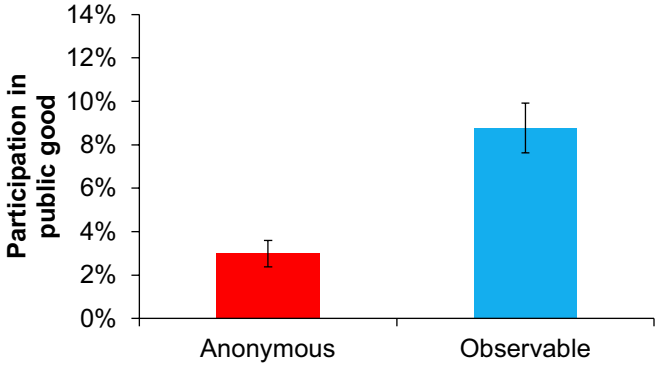



**Fig. 2.** We solicited 1,408 customers of a major electric utility for participation in a program designed to prevent blackouts. Residents signed up for the program on sheets posted in a communal area near their home, usually by a shared mailbox kiosk. We varied whether residents' neighbors could tell who signed up for the program: publicly posted sheets required residents to print their name and unit number (observable treatment) or only a code that does not reveal their identity (anonymous treatment). Observability tripled participation in the program.

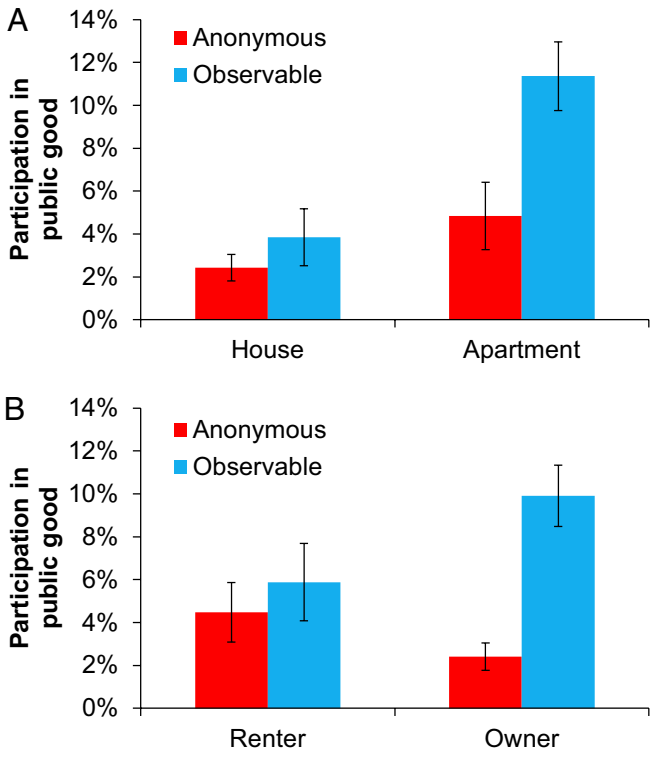



**Fig. 3.** Observability increased participation more in settings where reputational concerns matter more. (*A*) Observability increased participation more in apartment buildings where residents are more likely to interact with their neighbors in public spaces and sign-up sheets were typically posted in especially conspicuous locations, compared with row houses or individual homes, where neighbors are less likely to interact and sign-up sheets were less easily visible by others. (*B*) Similarly, observability increased participation more among those who own their homes/apartments relative to those who rent because renters are more transient and therefore likely to be less invested in long-term relationships with their neighbors.

Second, we test whether observability had a larger effect among those who own their homes/apartments relative to those who rent. Renters are more transient and therefore likely to be less invested in relationships with their neighbors. Thus, indirect reciprocity theory predicts that observability will have a larger effect among owners. As shown in Fig. 3*B*, the results were again consistent with this prediction: observability dramatically increased participation among owners (fraction of residents participating: anonymous = 0.024, observable = 0.099, $P < 0.01$, $n = 1015$) but had little effect on renters (fraction of residents participating: anonymous = 0.045, observable = 0.059, not significant, $n = 393$; yielding an estimated interaction of 0.046, $P < 0.01$).

Residents of apartment buildings and individual homes differ on any number of dimensions, as do those who own versus rent their homes. Thus, although the results are consistent with the predictions of the indirect reciprocity framework, alternative explanations of these results are possible. To partially address this issue, we ran the same analysis with additional controls such as Spanish language preference, ethnicity, missed payments, rebate use, and historical electrical use, and found even stronger results (*Supporting Information*). Therefore, differences on these dimensions do not account for the differential effects of observability seen in Figs. 2 and 3.

Finally, we provide evidence that the effect of observability is unique to public goods. Not participating in the program should carry the threat of social sanctions only if participation is considered

to be a public good. Therefore, indirect reciprocity theory predicts that observability should not increase participation among subjects who do not think of participation as a public good. To test this prediction, an additional 1,005 subjects received exactly the same treatment as described above, except that the mailers they received were stripped of any language that framed the demand response program as a public good. Consistent with our hypothesis, Fig. 4 shows that the effect of observability was reduced in this cohort (fraction of residents participating: anonymous = 0.061, observable = 0.086, not significant, $n = 1005$; estimated interaction between observability and the public good message in a pooled regression is 0.035, $P = 0.098$).

## Discussion

We have shown that indirect reciprocity promotes cooperation in a real-world public goods game affecting thousands of people. Making participation in the public good observable substantially increased sign-ups and did so significantly more than offering a cash incentive. Moreover, the effect of observability was larger in settings where individuals were more likely to have future interactions with those who observed them, and when participation was framed as a public good. These results provide evidence that reputational concerns were the driving force behind the effect of observability in our study.

Our study is part of a nascent literature exploring reputation and prosociality using field experiments. Consistent with our findings in the domain of energy efficiency, there is evidence that publicizing the names of donors increases the frequency of blood donation (70) as well as the level of giving to a college charity (71). Nonfinancial incentives involving reputation have also been shown to outperform monetary incentives in motivating the sale of condoms on behalf of a health organization in Nambia (72). Our work adds to these studies by directly manipulating observability, allowing a comparison with monetary incentives while avoiding other potential confounds present in previous experiments. We also test specific hypotheses generated by indirect reciprocity theory regarding when observability will and will not increase cooperation. Taken together, this body of work provides clear evidence that reputational incentives can be a powerful force for increasing cooperation in the field. Our paper in particular adds to efforts aimed at promoting energy conservation via nonfinancial incentives, such as providing people information about their own energy use and how it compares with the energy use of their neighbors (79–82).

A question arising from our study is the extent to which our subjects were conscious of their indirect reciprocity motives. One possibility is that they explicitly considered the reputational costs of not participating in the observable public goods treatment. Alternatively, they may have learned or evolved sensitivity to subtle cues that subconsciously increased their desire to participate when their decisions were observable, as has been shown in other settings (62, 63, 65, 83). Perhaps the degree of "warm glow" they feel is sensitive to the degree of observability in their environment and the likelihood of interacting with observers in the future. Subsequent studies should further investigate this issue.

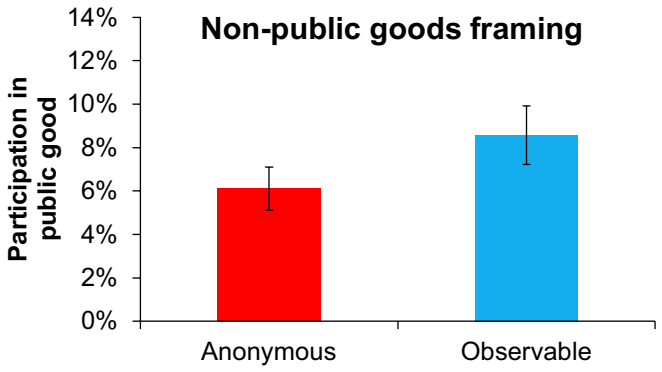



**Fig. 4.** We solicited an additional 1,005 customers with exactly the same treatment as described above, except that the informational materials they received were stripped of any language that framed the blackout prevention program as a public good. The effect of observability was dramatically reduced among subjects who did not receive the public good framing.

A related issue is the universality of reputation concerns. Observability can promote cooperation, but only in populations where the proper social norms are in place. For example, in a laboratory experiment in the United States, making public goods contributions observable by linking the PGG to a set of pairwise Prisoner's Dilemma games led to high contributions (53). However, when the same experiment was run using students in Romania, no such positive effects were observed because the Romanians did not sanction bad behavior in the PGG (84). Similarly, providing feedback on how one's energy use compares with one's neighbors had reduced consumption among American liberals but may have had the opposite effect among conservatives (80). Studying the interaction between norms and institutional policies is an important direction for future research.

In our experiment, the observability mechanism was designed so that participation was automatically displayed to all: because sign-up sheets were posted in public areas, no special effort was required by individuals to spread reputational information. Most indirect reciprocity models, however, rely on individuals communicating information about the observed actions of others (21). Fortunately, we are more than happy to talk about how others have behaved: gossip is a central element of human communication (85, 86). However, why did we come to have this predilection for gossiping about the previous behavior of others? Why spend time and effort on evaluating others, and why give honest evaluations of competitors? Indirect reciprocity itself offers a potential answer: providing honest information or not is another game of cooperation and defection, which is also linked to reputation. Your reputation can be damaged not just by defection in the primary public goods game, but also by the distribution of incomplete or false information. Another important question involves large-scale reputation systems such as those used by the online market eBay (87) or the business rating website Yelp (88): to what extent does our intrinsic desire to gossip extend to these more distributed settings? Why do people bother to leave evaluations, and how can secondary reputation systems be designed to encourage honest feedback? Exploring these issues is an important direction for further study.

Indirect reciprocity offers a powerful tool for promoting cooperation in contexts of great societal importance. Here, we offer quantitative evidence for one example: curbing electricity use during periods of high demand. However, this is just one of many such opportunities (70–72, 89). For example, people might be induced to drive more efficient cars if all vehicles bore a visible indication of fuel efficiency, perhaps via mandated color coding of license plates for the most efficient and most wasteful vehicles. Or home energy use might be reduced if utility companies made individuals' power use statistics publicly available. One might even apply this logic to scientific discovery: a measure of "scientific carbon efficiency" could by calculated by dividing an author's number of citations (or h-index) by the number of miles flown to attend conferences. Of course, privacy is an important issue that must be balanced against the benefits of reputational pressure. However, there are also indirect reciprocity applications that do not infringe on the privacy rights of individuals. For example, businesses might reduce their environmental impact if they were required to disclose the overall carbon footprint of their operations. Reputational concerns might discourage financial institutions from taking excessive risk because of changes in

the price at which they have to borrow. Or government agencies might reduce waste if the relevant statistics were readily accessible to the public. Developing interventions that harness indirect reciprocity is a promising direction for future public policy initiatives.

## Methods

We administered the field experiment with the collaboration of the Pacific Gas and Electric Company (PG&E), a regulated utility that operates the majority of Northern California's retail residential electricity market. The experiment was incorporated into a routine marketing effort for a demand response program called SmartAC, which is designed to help prevent or shorten power interruptions by curbing demand from central air conditioners on days with unusually high demand, or in the case of an unexpected plant or transmission failure. The program is voluntary; subjects who participate contribute to a public good by contributing to the stability of the electrical grid in all of California, at the cost of some personal inconvenience and possibly some discomfort. The SmartAC switch is installed free-of-charge. At the time of the program, participants received a $25 check for signing up. The SmartAC program is a typical demand-side management, direct load control, or load shedding program.

Subjects in the field experiment were residential customers living in homeowners associations (HOAs) and one rental complex in Santa Clara County. We focused on tenants of HOAs because it was necessary to choose residences with public spaces where sign-up sheets could be posted. We focused on Santa Clara County because PG&E had not marketed in this area before the field experiment. Furthermore, Santa Clara County is hot enough that customers there were likely to have air conditioners, and dense enough to have a sufficient number of HOAs. Finally, we restricted the analysis to HOAs where all residents were known to have central air conditioning because central air conditioning was required to participate in the SmartAC program.

We invited subjects to participate in the program by sliding marketing materials under subjects' doors, placing them on their doorstep, or mailing materials to subjects. The materials included an informational letter describing the program and an instruction card that directed subjects to sign up for the program on sign-up sheets posted next to their mailboxes or in another central location. We left the sign-up sheets up for 3 to 10 d, depending on managers' preferences, the weather, and other conditions. After distributing the marketing materials, we removed the sign-up sheets, noted subjects' participation decisions, and provided the list of participants to PG&E's contractor for processing and installation. Note that subjects were not aware that they were participating in an experiment. This study is therefore classified as a natural field experiment (90).

In the experiment's main treatment, we varied observability by varying the design of the sign-up sheets on which subjects register for the program: some sheets were designed so that subjects' identities were easily revealed to others who observed the sign-up sheet whereas others were designed to conceal subjects' identities. In the latter "anonymous'" design, the fields for subjects' names and apartment numbers were omitted from the sign-up sheet. Instead, subjects were identified only by their randomly generated personal code.

Simultaneously, we varied the design of the marketing materials along two dimensions. First, we varied whether the materials framed the decision to sign up as a contribution to a public good that would benefit others, or just as a new feature being offered by PG&E. Second, we varied whether subjects were offered a $25 incentive for signing up for the program. See *Supporting Information* for further details of the experimental design.

**ACKNOWLEDGMENTS.** This research was partially funded by National Science Foundation Proposal 0752818, Russell Sage Foundation Award 980803, the Stigler Center for the Study of the Economy and the State's Energy Initiative, the University of Chicago Graduate School of Business, and the John Templeton Foundation. Any opinions expressed in this article are those of the authors and not of the Federal Trade Commission or any individual Commissioner.

1. Hardin G (1968) The tragedy of the commons. The population problem has no technical solution; it requires a fundamental extension in morality. *Science* 162(3859): 1243–1248.
2. Ostrom E (1990) *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge Univ Press, Cambridge, UK).
3. Axelrod R (1984) *The Evolution of Cooperation* (Basic Books, New York).
4. Nowak MA, Highfield R (2011) *SuperCooperators: Altruism, Evolution, and Why We Need Each Other to Succeed* (Free Press, New York).
5. Churchland PS (2012) *Braintrust: What Neuroscience Tells Us About Morality* (Princeton Univ Press, Princeton).
6. Oakley B, Knafo A, Madhavan G, Wilson DS, eds (2011) *Pathological Altruism* (Oxford Univ Press, Oxford).
7. Levin SA (2006) Learning to live in a global commons: Socioeconomic challenges for a sustainable environment. *Ecol Res* 21(3):328–333.
8. Walker B, et al. (2009) Environment. Looming global-scale failures and missing institutions. *Science* 325(5946):1345–1346.
9. Helbing D, Yu W (2009) The outbreak of cooperation among success-driven individuals under noisy conditions. *Proc Natl Acad Sci USA* 106(10):3680–3685.
10. Sigmund K (2010) *The Calculus of Selfishness* (Princeton Univ Press, Princeton).
11. Nowak MA (2006) Five rules for the evolution of cooperation. *Science* 314(5805): 1560–1563.
12. Nowak MA, Tarnita CE, Antal T (2010) Evolutionary dynamics in structured populations. *Philos Trans R Soc Lond B Biol Sci* 365(1537):19–30.
13. Dasgupta P (2012) Dark matters: Exploitation as cooperation. *Journal of Theoretical Biology* 299:180–187.
14. Choi JK, Bowles S (2007) The coevolution of parochial altruism and war. *Science* 318(5850):636–640.
15. Carlton DW, Perloff JM (1994) *Modem Industrial Organization* (HarperCollins, New York).
16. Cela-Conde CJ, Ayala FJ (2007) *Human Evolution: Trails from the Past* (Oxford University Press, Oxford).
17. Fu F, Hauert C, Nowak MA, Wang L (2008) Reputation-based partner choice promotes cooperation in social networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 78(2 Pt 2): 026117.
18. Manapat ML, Nowak MA, Rand DG (2012) Information, irrationality and the evolution of trust. *J Econ Behav Organ*, in press.
19. Nowak MA, Page KM, Sigmund K (2000) Fairness versus reason in the ultimatum game. *Science* 289(5485):1773–1775.
20. Nowak MA, Sigmund K (1998) Evolution of indirect reciprocity by image scoring. *Nature* 393(6685):573–577.
21. Nowak MA, Sigmund K (2005) Evolution of indirect reciprocity. *Nature* 437(7063): 1291–1298.
22. Ohtsuki H, Iwasa Y, Nowak MA (2009) Indirect reciprocity provides only a narrow margin of efficiency for costly punishment. *Nature* 457(7225):79–82.
23. Masuda N, Ohtsuki H (2007) Tag-based indirect reciprocity by incomplete social information. *Proc Biol Sci* 274(1610):689–695.
24. Ohtsuki H, Iwasa Y (2006) The leading eight: Social norms that can maintain cooperation by indirect reciprocity. *J Theor Biol* 239(4):435–444.
25. Panchanathan K, Boyd R (2004) Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature* 432(7016):499–502.
26. Kandori M (1992) Social norms and community enforcement. *Rev Econ Stud* 59:63–80.
27. Brandt H, Sigmund K (2006) The good, the bad and the discriminator—errors in direct and indirect reciprocity. *J Theor Biol* 239(2):183–194.
28. Pacheco JM, Santos FC, Chalub FACC (2006) Stern-judging: A simple, successful norm which promotes cooperation under indirect reciprocity. *PLOS Comput Biol* 2(12):e178.
29. Ohtsuki H, Iwasa Y (2007) Global analyses of evolutionary dynamics and exhaustive search for social norms that maintain cooperation by reputation. *J Theor Biol* 244(3): 518–531.
30. Leimar O, Hammerstein P (2001) Evolution of cooperation through indirect reciprocity. *Proc Biol Sci* 268(1468):745–753.
31. Panchanathan K, Boyd R (2003) A tale of two defectors: The importance of standing for evolution of indirect reciprocity. *J Theor Biol* 224(1):115–126.
32. Nowak MA, Sigmund K (1998) The dynamics of indirect reciprocity. *J Theor Biol* 194(4):561–574.
33. Suzuki S, Akiyama E (2007) Evolution of indirect reciprocity in groups of various sizes and comparison with direct reciprocity. *J Theor Biol* 245(3):539–552.
34. Ohtsuki H, Iwasa Y (2004) How should we define goodness?—reputation dynamics in indirect reciprocity. *J Theor Biol* 231(1):107–120.
35. Manapat ML & Rand DG (2012) Delayed and inconsistent information and the evolution of trust. *Dyn Games Appl* 2:401–410.
36. Uchida S, Sigmund K (2010) The competition of assessment rules for indirect reciprocity. *J Theor Biol* 263(1):13–19.
37. Nakamura M, Masuda N (2011) Indirect reciprocity under incomplete observation. *PLOS Comput Biol* 7(7):e1002113.
38. Suzuki S, Akiyama E (2007) Three-person game facilitates indirect reciprocity under image scoring. *J Theor Biol* 249(1):93–100.
39. Berger U (2011) Learning to cooperate via indirect reciprocity. *Games Econ Behav* 72(1):30–37.
40. Mani A, Rahwan I, Pentland A (2013) Inducing Peer Pressure to Promote Cooperation. *Scientific Reports* 3:1735.
41. Seinen I, Schram A (2006) Social status and group norms: Indirect reciprocity in a repeated helping experiment. *Eur Econ Rev* 50(3):581–602.
42. Ule A, Schram A, Riedl A, Cason TN (2009) Indirect punishment and generosity toward strangers. *Science* 326(5960):1701–1704.
43. Milinski M, Semmann D, Bakker TCM, Krambeck HJ (2001) Cooperation through indirect reciprocity: Image scoring or standing strategy? *Proc Biol Sci* 268(1484): 2495–2501.
44. Milinski M, Semmann D, Krambeck HJ (2002) Donors to charity gain in both indirect reciprocity and political reputation. *Proc Biol Sci* 269(1494):881–883.
45. Milinski M, Semmann D, Krambeck HJ (2002) Reputation helps solve the 'tragedy of the commons'. *Nature* 415(6870):424–426.

46. Rockenbach B, Milinski M (2006) The efficient interaction of indirect reciprocity and costly punishment. *Nature* 444(7120):718–723.
47. Semmann D, Krambeck H-J, Milinski M (2005) Reputation is valuable within and outside one's own social group. *Behav Ecol Sociobiol* 57(6):611–616.
48. Wedekind C, Milinski M (2000) Cooperation through image scoring in humans. *Science* 288(5467):850–852.
49. Pfeiffer T, Tran L, Krumme C, Rand DG (2012) The value of reputation. *J R Soc Interface* 9(76):2791–2797.
50. Wedekind C, Braithwaite VA (2002) The long-term benefits of human generosity in indirect reciprocity. *Curr Biol* 12(12):1012–1015.
51. Bolton GE, Katok E, Ockenfels A (2005) Cooperation among strangers with limited information about reputation. *J Public Econ* 89(8):1457–1468.
52. Jacquet J, Hauert C, Traulsen A, Milinski M (2011) Shame and honour drive cooperation. *Biol Lett* 7(6):899–901.
53. Rand DG, Dreber A, Ellingsen T, Fudenberg D, Nowak MA (2009) Positive interactions promote public cooperation. *Science* 325(5945):1272–1275.
54. Choi J-K, Ahn TK (2013) Strategic reward and altruistic punishment support cooperation in a public goods game experiment. *J Econ Psychol*, 35:17–30.
55. Rege M, Telle K (2004) The impact of social approval and framing on cooperation in public good situations. *J Public Econ* 88(7):1625–1644.
56. Andreoni J, Petrie R (2004) Public goods experiments without confidentiality: A glimpse into fund-raising. *J Public Econ* 88(7):1605–1623.
57. Tadelis S (2007) The power of shame and the rationality of trust. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1006169. Accessed April 20, 2013.
58. Andreoni J, Bernheim BD (2009) Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica* 77(5):1607–1636.
59. Linardi S, McConnell MA (2008) Volunteering and Image Concerns. *California Institute of Technology Working Paper*.
60. List JA, Berrens RP, Bohara AK, Kerkvliet J (2004) Examining the role of social isolation on stated preferences. *Am Econ Rev* 94(3):741–752.
61. Kossmeier S, Ariely D, Bracha A (2009) Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *Am Econ Rev* 99(1):544–555.
62. Haley KJ, Fessler DMT (2005) Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evol Hum Behav* 26:245–256.
63. Burnham T, Hare B (2007) Engineering Human Cooperation. *Hum Nat* 18(2):88–108.
64. Powell KL, Roberts G, Nettle D (2012) Eye Images Increase Charitable Donations: Evidence From an Opportunistic Field Experiment in a Supermarket. *Ethology* 118(11):1096–1101.
65. Ernest-Jones M, Nettle D, Bateson M (2011) Effects of eye images on everyday cooperative behavior: a field experiment. *Evol Hum Behav* 32(3):172–178.
66. Rand DG, Greene JD, Nowak MA (2012) Spontaneous giving and calculated greed. *Nature* 489(7416):427–430.
67. Schulz JF, Fischbacher U, Thöni C, & Utikal V (In press) Affect and fairness: Dictator games under cognitive load. *Journal of Economic Psychology* (0).
68. Rand DG, et al. (2013) Intuitive Cooperation and the Social Heuristics Hypothesis: Evidence from 15 Time Constraint Studies. *Available at SSRN: http://ssrn.com/abstract=2222683*.
69. Cornelissen G, Dewitte S, Warlop L (2011) Are social value orientations expressed automatically? Decision making in the dictator game. *Pers Soc Psychol Bull* 37(8):1080–1090.
70. Lacetera N, Macis M (2010) Social image concerns and prosocial behavior: Field evidence from a nonlinear incentive scheme. *J Econ Behav Organ* 76(2):225–237.
71. Karlan D, McConnell MA (2012) *Hey Look at Me: The Effect of Giving Circles on Giving* (National Bureau of Economic Research, Cambridge, MA).
72. Ashraf N, Bandiera O, Jack K (2012) No margin, no mission? A field experiment on incentives for pro-social tasks. *Harvard Business School Working Papers*. Available at www.hbs.edu/faculty/Pages/download.aspx?name=12-008.pdf. Accessed April 20, 2013.
73. Levitt SD, List JA (2007) What do laboratory experiments measuring social preferences reveal about the real world? *J Econ Perspect* 21:153–174.
74. US Committee on Technology (2011) *A Policy Framework for the 21st Century Grid: Enabling Our Secure Energy Future* (Executive Office of the President: National Science and Technology Council, Washington).
75. Chupka MW, Earle R, Fox-Penner P, Hledik R (2008) *Transforming America's Power Industry: The Investment Challenge 2010–2030* (The Edison Foundation, Washington).
76. Kathan D, et al. (2011) *Assessment of Demand Response and Advanced Metering: Staff Report* (Federal Energy Regulatory Commission, Washington).
77. Gneezy U, Rustichini A (2000) Pay Enough or Don't Pay At All. *Q J Econ* 115(3):791–810.
78. Bertrand M, Karlan D, Mullainathan S, Shafir E, Zinman J (2010) What's advertising content worth? Evidence from a consumer credit marketing field experiment. *Q J Econ* 125(1):263–306.
79. Ayres I, Raseman S, Shih A (2009) *Evidence from Two Large Field Experiments That Peer Comparison Feedback Can Reduce Residential Energy Usage* (National Bureau of Economic Research, Cambridge, MA).
80. Costa DL, Kahn ME (2010) *Energy Conservation" Nudges" and Environmentalist Ideology: Evidence from a Randomized Residential Electricity Field Experiment* (National Bureau of Economic Research, Cambridge, MA).
81. Thompson C (2007) Clive Thompson thinks: Desktop orb could reform energy hogs. *Wired Magazine* 15(8).
82. Minosi A, et al. (2003) Intelligent, low-power and low-cost measurement system for energy consumption. *IEEE International Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems* (IEEE, Washington), pp 125–130.
83. Bateson M, Nettle D, Roberts G (2006) Cues of being watched enhance cooperation in a real-world setting. *Biol Lett* 2(3):412–414.
84. Ellingsen T, Herrmann B, Nowak MA, Rand DG, Tarnita CE (2012) Civic Capital in Two Cultures: The Nature of Cooperation in Romania and USA. Available at http://ssrn.com/abstract=2179575. Accessed April 20, 2013.
85. Dunbar RIM, Marriott A, Duncan NDC (1997) Human conversational behavior. *Hum Nat* 8(3):231–246.
86. Sommerfeld RD, Krambeck H-J, Semmann D, Milinski M (2007) Gossip as an alternative for direct observation in games of indirect reciprocity. *Proc Natl Acad Sci USA* 104(44):17435–17440.
87. Resnick P, Zeckhauser R, Swanson J, Lockwood K (2006) The value of reputation on eBay: A controlled experiment. *Exp Econ* 9(2):79–101.
88. Luca M (2011) Reviews, reputation, and revenue: The case of Yelp.com. Available at http://hbswk.hbs.edu/item/6833.html. Accessed April 20, 2013.
89. Rand DG, Nowak MA (2009) Name and Shame. *New Sci* 204(2734):28–29.
90. Harrison GW, List JA (2004) Field experiments. *J Econ Lit* 42(4):1009–1055.

# Potential follow-up increases private contributions to public goods

Todd Rogers[a,1], John Ternovski[a], and Erez Yoeli[b]

[a]Harvard Kennedy School, Harvard University, Cambridge, MA 02138; and [b]Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved March 18, 2016 (received for review December 16, 2015)

**People contribute more to public goods when their contributions are made more observable to others. We report an intervention that subtly increases the observability of public goods contributions when people are solicited privately and impersonally (e.g., mail, email, social media). This intervention is tested in a large-scale field experiment ($n$ = 770,946) in which people are encouraged to vote through get-out-the-vote letters. We vary whether the letters include the message, "We may call you after the election to ask about your voting experience." Increasing the perceived observability of whether people vote by including that message increased the impact of the get-out-the-vote letters by more than the entire effect of a typical get-out-the-vote letter. This technique for increasing perceived observability can be replicated whenever public goods solicitations are made in private.**

public goods | reputation | observability | get-out-the-vote | field experiment

How can we increase contributions to public goods—to get donors to give more to charity, citizens to vote, households to consume less energy, drivers to carpool, and patients to take all of their antibiotics? One of the best ways is to make contributions more observable (1, 2), as demonstrated by a large body of laboratory experiments (3–9) and a growing body of field experiments (for a review, see ref. 2) in a variety of settings, including energy conservation (10), blood donations (11), national park contributions (12), and voting (13).

Observability increases contributions to public goods such as voting or charitable giving because observability allows contributions to affect reputations. Individuals who are observed to have contributed can be held in good standing and rewarded in subsequent relationships, either when others are more likely to engage them in a relationship in the first place (this is called partner choice; e.g., refs. 14 and 15) or when others are more cooperative with them during an existing relationship (this is called indirect reciprocity; e.g., refs. 16–21). And, individuals who are observed to not contribute can be held in poor standing.

Even subtle cues of observability can increase contributions. In fact, observability can affect contributions when the reputational consequences of one's choice have been entirely eliminated (22, 23). An example is eyespots: simply displaying a picture of a face or an abstraction resembling a face increases contributions (24, 25). Such effects imply that the psychology governing our reputations operates at the intuitive level (24)—that is, people do not necessarily deliberate over the reputational gains of every cooperative action, and instead rely on heuristics. Such an intuitive psychology might develop if the heuristics usually work (26, 27). For example, if seeing something that looks like a face is usually an accurate indication that someone is watching, then it may pay to give more whenever in the presence of something that looks like a face, even though a clever researcher may exploit this heuristic to induce people into giving a little more in an experiment. Moreover, there are reputational gains to cooperating without deliberating about the decision. Namely, people are perceived as being more trustworthy when cooperation is the automatic behavior. This, too, can lead people to rely more on cues and heuristics (28).

In this paper, we report the results of a large field experiment in which we subtly increased perceived observability to motivate contributions to a real-world public good: voting. The experiment involved sending get-out-the-vote (GOTV) letters to citizens before the 2010 General Election (total $n$ = 770,946). There were three conditions. Those assigned to the best practices condition ("Best"; $n$ = 346,929) were mailed a GOTV letter containing several messaging elements that have been shown to increase turnout (see Fig. S1 for complete reproduction of all letters and Tables S1 and S2 for balance checks across treatments). Those assigned to the best practices-plus-increased probability of observability condition ("Best plus Observable"; $n$ = 347,054) were sent a GOTV letter that was identical to the one sent to those in the Best condition with two exceptions. First, at the top right corner of the page, these letters included the message, "You may be called after the election to discuss your experience at the polls." Second, a paragraph was also added at the end of the GOTV letter reinforcing this message. See Fig. 1. Those assigned to the control received no GOTV letter ($n$ = 76,963).

The outcome measure of interest is voter turnout. Those assigned to the Best condition voted at meaningfully higher rates than those assigned to control (41.36% vs. 40.88%, $z$ = 2.52, $P$ = 0.012). All analyses control for preexperiment stratifications, although results hold without these controls. The GOTV letter sent to those in the Best condition increased turnout by 0.48 percentage points. Meta-analyses of 79 experiments examining the impact of typical nonpartisan GOTV letters show that the average treatment effect is 0.194 percentage points (29). This means that the GOTV letter sent to those in the Best condition was more than twice as effective as the typical GOTV letter [$F_{(1,770915)}$ = 2.37, $P$ = 0.12]. As described in *Methods*, one reason this GOTV letter may have been especially

## Significance

**Reputational concern is one reason people perform behaviors that are good for society but have little benefit for individuals (e.g., energy efficiency, donation, recycling, voting). In order for a behavior to influence reputations, it must be observable. However, many strategies for encouraging these behaviors involve communicating privately and impersonally (e.g., mail, email, social media) with little or no observability. We report a large-scale field experiment ($N$ = 770,946) examining a technique for harnessing the benefits of observability when encouraging these behaviors privately. Get-out-the-vote letters become substantially more effective when they say, "We may call you after the election to ask about your voting experience." This technique can be widely used to encourage society-benefiting behaviors.**

Author contributions: T.R. and J.T. designed research; T.R. and J.T. performed research; T.R. and J.T. analyzed data; and T.R., J.T., and E.Y. wrote the paper.

The authors declare no conflict of interest.

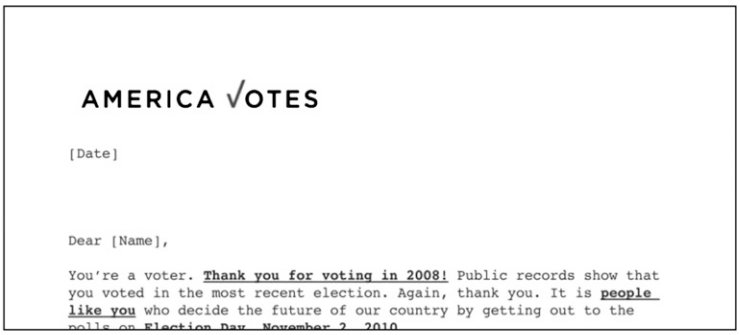
This article is a PNAS Direct Submission.

Data deposition: The data and analysis code have been deposited in the Open Science Framework's archive, https://osf.io/thxj5.

[1]To whom correspondence should be addressed. Email: Todd_Rogers@hks.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1524899113/-/DCSupplemental.

*Best Practices Condition*

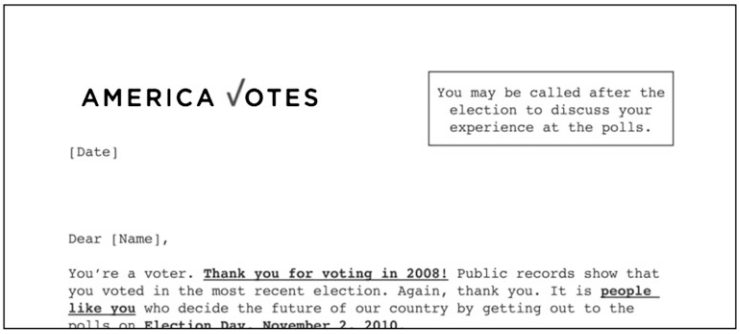

AMERICA VOTES

[Date]

Dear [Name],

You're a voter. **Thank you for voting in 2008!** Public records show that you voted in the most recent election. Again, thank you. It is **people like you** who decide the future of our country by getting out to the polls on **Election Day, November 2, 2010**.

*Best Practices + Increased Probability of Observability Condition*


AMERICA VOTES

You may be called after the election to discuss your experience at the polls.

[Date]

Dear [Name],

You're a voter. **Thank you for voting in 2008!** Public records show that you voted in the most recent election. Again, thank you. It is **people like you** who decide the future of our country by getting out to the polls on **Election Day, November 2, 2010**.

*Letters are fully reproduced in S3. Note that letter in best practices + heightened observability condition ends with a paragraph reinforcing observability.*

**Fig. 1.** Treatment letters.

potent is that it already contained several elements highlighting the observability of whether one votes. The content of the GOTV letter sent to those in the Best-plus-Observable condition amplified the suggested observability above and beyond what was suggested in the letter sent to those in the Best condition. At \$0.34 per letter, the letter sent to those in the Best condition cost \$71 per net vote.

The GOTV letter sent to those assigned to the Best-plus-Observable condition increased turnout by 0.72 percentage points compared with the control group (41.60% vs. 40.88%, $z = 3.80$, $P < 0.001$). This GOTV letter was more effective than that sent to those in the Best condition (41.60% vs. 41.36%, $z = 2.13$, $P = 0.033$). That is, adding observability to the already-effective GOTV letter sent to those in the Best condition increased the impact of the GOTV letter by 0.22 percentage points—a 51% improvement that is larger than the average impact of the typical GOTV letter. The GOTV letter sent to those in the Best-plus-Observable condition was more than three times as effective as the typical GOTV letter [$F_{(1,770915)} = 7.95$, $P < 0.01$]. The GOTV letter sent to those in the Best-plus-Observable condition cost \$47 per net vote—less than one-third of the \$175 per net vote generated from the typical GOTV letter (29). See Fig. 2.

Why does the prospect of a follow-up call increase voting? As with most studies of observability, we cannot rule out that people consciously responded to the intervention—that they deliberated on the benefits of voting and evaluated them as greater in the Best-plus-Observable condition. However, the fact that the future call was uncertain, and that if it did happen it would entail a conversation with a total stranger, suggests this is unlikely. Instead, it seems more plausible that the intervention acted on a nonconscious level, as in the eyespot studies (24, 25). For example, the prospect of a follow-up might have activated feelings of accountability (30) or served as a reminder of future social interactions in which voting might be discussed (e.g., mental simulation).

We also cannot rule out that factors not directly related to public goods contributed to the intervention's success. In particular, the intervention might motivate one to vote simply to avoid disappointing or confronting a concerned party. Or people may vote to avoid the unpleasant experience of having to lie—to claim that one voted when one did not (31). If so, the intervention may work in additional settings. For example, a counselor or advisor may be able to motivate students to follow through on their assignments and studying by scheduling weekly meetings.

We speculate that repeated attempts to increase people's perceptions of observability by suggesting the prospect of a follow-up contact will become decreasingly effective if follow-up contacts are not made. The intervention will lose credibility, and so it will not heighten perceptions of the reputation consequences of contributing. Therefore, we suggest that this intervention will be most effective when the chance of follow-up is credible (e.g., because a follow-up survey or in-person interaction is already planned).

Our study makes three contributions. The first is practical. Many solicitations for contributions are made privately—for example, by mail, over email, or by posting on social media platforms such as Facebook and Twitter. These account for a large portion of fundraising: direct mail fundraising accounts for roughly 20% of all charitable donations, and online fundraising accounts for another 7% and is rapidly growing (32). Candidates and political groups regularly encourage constituents to vote using these same private communications media. Thus, for many practitioners, our results provide a practical, inexpensive, and effective strategy for increasing observability when soliciting public goods contributions via private communication.

Second, our results add to the field evidence that public goods contributions can be increased by making contributions more observable—even by merely suggesting that there may be magnified observability. Finally, our results provide additional evidence that

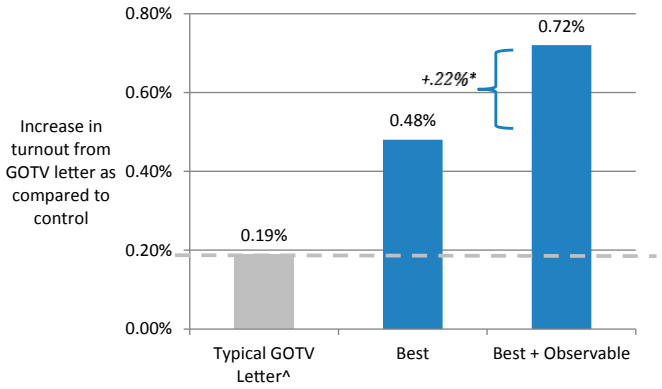



**Fig. 2.** Adding observability to a GOTV letter increases impact more than the average impact of a typical GOTV letter. ^From metaanalysis of 79 randomized experiments of typical GOTV letters (29); turnout in the control group was 40.88%; *$P < 0.05$; Best and Best-plus-Observable bars represent comparison to this experiment's control group, whereas typical GOTV letter bar represents comparison to control groups included in the metaanalysis.

voting can be increased by interventions that might affect reputations (e.g., refs. 13 and 33–37).

## Methods

The GOTV letters were sent from an independent 501(c)(4) organization that was likely unfamiliar to recipients, America Votes. America Votes selected the experiment universe based on three criteria using data provided by the political data vendor Catalist, LLC (38). First, individuals had to reside in 1 of 29 targeted battleground congressional districts chosen based on the organization's political objectives and expectation that the elections would be close. Second, only one individual per household could be included. Third, using predictive models developed by Catalist, individuals had to be predicted to be politically "progressive" and to have a low-to-moderate propensity to vote in the 2010 General Election. This resulted in a population that was 60% female, 15% African American, and averaged 43 y of age. The experiment universe included 645,035 individuals who voted in the 2008 General Election and 160,721 individuals who did not vote in the 2008 General Election but who had registered to vote in the 2010 General Election. Before being randomly assigned to one of the three conditions, the experiment universe was stratified by whether individuals voted in the 2008 election, and by their congressional district.

The GOTV letters emphasized the descriptive social norm that many others would vote (33). They also reinforced the civic identity by highlighting that "people like you" will vote (36). Another messaging element in the GOTV letters involved a callout box in which targets were to write their voting plans, reflecting work on the power of implementation intentions on turnout and other health behaviors (34, 39). The GOTV letters also expressed gratitude for the targets' past political actions, and a hope that public records would show that targets will have voted in the upcoming 2010 election (35). For those who had voted in the 2008 General Election, the letter thanked them for voting in 2008, and for those who had not voted in the 2008 General Election but had registered to vote in the 2010 General Election, it thanked them for registering. This was the only difference in the messaging content between those who had voted in 2008 and those who had not. Note that this GOTV letter already indicates to voters that their behavior is observable. This indication could mute any effect of adding an explicit suggestion that whether people vote may be observable.

Voter turnout data were collected by Catalist, LLC, from publicly reported administrative records. Turnout data were missing for one district, MA-10, which has been excluded from all analyses. We administered a survey to a subsample of targets 2 mo after the election, but it is not relevant to this manuscript. All analyses use logistic regression controlling for preexperiment strata: dummies for congressional districts, and 2008 vote history.

**ACKNOWLEDGMENTS.** We thank America Votes, Analyst Institute, and Catalist, LLC, for collaboration and data.

1. Rand DG, Yoeli E, Hoffman M (2014) Harnessing reciprocity to promote cooperation and the provisioning of public goods. *Policy Insights Behav Brain Sci* 1(1):263–269.
2. Kraft-Todd GT, Norton MI, Rand DG (2015) The selfishness of selfless people: Harnessing the theory of self-concept maintenance to increase charitable giving. Available at ssrn.com/abstract=2568869. Accessed April 3, 2016.
3. Andreoni J, Petrie R (2004) Public goods experiments without confidentiality: A glimpse into fund-raising. *J Public Econ* 88:1605–1623.
4. List JA, Berrens RP, Bohara AK, Kerkvliet J (2004) Examining the role of social isolation on stated preferences. *Am Econ Rev* 94:741–752.
5. Rege M, Telle K (2004) The impact of social approval and framing on cooperation in public good situations. *J Public Econ* 88:1625–1644.
6. Tadelis S (2007) The power of shame and the rationality of trust. Available at papers.ssrn.com/sol3/papers.cfm?abstract_id=1006169. Accessed April 3, 2016.
7. Andreoni J, Bernheim BD (2009) Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica* 77:1607–1636.
8. Linardi S, McConnell MA (2008) Volunteering and image concerns. Available at people.hss.caltech.edu/~slinardi/papers/Volunteer_Jan08.pdf. Accessed April 3, 2016.
9. Jacquet J, Hauert C, Traulsen A, Milinski M (2011) Shame and honour drive cooperation. *Biol Lett* 11:rsbl20110367.
10. Yoeli E, Hoffman M, Rand DG, Nowak MA (2013) Powering up with indirect reciprocity in a large-scale field experiment. *Proc Natl Acad Sci USA* 110(Suppl 2):10424–10429.
11. Lacetera N, Macis M (2010) Social image concerns and prosocial behavior: Field evidence from a nonlinear incentive scheme. *J Econ Behav Organ* 76:225–237.
12. Bandiera O, Barankay I, Rasul I (2005) Social preferences and the response to incentives: Evidence from personnel data. *Q J Econ* 120:917–962.
13. Gerber AS, Green DP, Larimer CW (2008) Social pressure and voter turnout: Evidence from a large-scale field experiment. *Am Polit Sci Rev* 102(01):33–48.
14. Glazer A, Konrad KA (1996) A signaling explanation for charity. *Am Econ Rev* 86:1019–1028.
15. Gintis H, Smith EA, Bowles S (2001) Costly signaling and cooperation. *J Theor Biol* 213(1):103–119.
16. Milinski M, Semmann D, Krambeck H-J (2002) Reputation helps solve the "tragedy of the commons." *Nature* 415(6870):424–426.
17. Panchanathan K, Boyd R (2004) Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature* 432(7016):499–502.
18. Nowak MA, Sigmund K (2005) Evolution of indirect reciprocity. *Nature* 437(7063):1291–1298.
19. Rand DG, Dreber A, Ellingsen T, Fudenberg D, Nowak MA (2009) Positive interactions promote public cooperation. *Science* 325(5945):1272–1275.
20. Balafoutas L, Nikiforakis N, Rockenbach B (2014) Direct and indirect punishment among strangers in the field. *Proc Natl Acad Sci USA* 111(45):15924–15927.
21. McNamara JM, Doodson P (2015) Reputation can enhance or suppress cooperation through positive feedback. *Nat Commun* 6:6134.
22. Dana J, Cain DM, Dawes RM (2006) What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organ Behav Hum Decis Process* 100:193–201.
23. Cain DM, Dana J, Newman GE (2014) Giving versus giving in. *Acad Management Ann* 8:505–533.
24. Haley KJ, Fessler DMT (2005) Nobody's watching? *Evol Hum Behav* 26:245–256.
25. Bateson M, Nettle D, Roberts G (2006) Cues of being watched enhance cooperation in a real-world setting. *Biol Lett* 2(3):412–414.
26. Bear A, Rand DG (2016) Intuition, deliberation, and the evolution of cooperation. *Proc Natl Acad Sci USA* 113(4):936–941.
27. Rand DG, et al. (2014) Social heuristics shape intuitive cooperation. *Nat Commun* 5:3677.
28. Hoffman M, Yoeli E, Nowak MA (2015) Cooperate without looking: Why we care what people think and not just what they do. *Proc Natl Acad Sci USA* 112(6):1727–1732.
29. Green DP, McGrath MC, Aronow PM (2013) Field experiments and the study of voter turnout. *J Elections Public Opin Parties* 23:27–48.
30. Lerner JS, Tetlock PE (1999) Accounting for the effects of accountability. *Psychol Bull* 125(2):255–275.
31. DellaVigna S, List JA, Malmendier U, Rao G (2014) Voting to tell others. National Bureau of Economic Research Working paper 19832. Available at www.nber.org/papers/w19832.
32. MacLaughlin S (2013) *Charitable Giving Report: How Nonprofit Fundraising Performed in 2012* (Blackbaud, Charleston, SC), pp 1–19.
33. Gerber AS, Rogers T (2009) Descriptive social norms and motivation to vote: Everybody's voting and so should you. *J Polit* 71(01):178–191.
34. Nickerson DW, Rogers T (2010) Do you have a voting plan?: Implementation intentions, voter turnout, and organic plan making. *Psychol Sci* 21(2):194–199.
35. Panagopoulos C (2011) Thank you for voting: Gratitude expression and voter mobilization. *J Polit* 73:707–717.
36. Bryan CJ, Walton GM, Rogers T, Dweck CS (2011) Motivating voter turnout by invoking the self. *Proc Natl Acad Sci USA* 108(31):12653–12656.
37. Bond RM, et al. (2012) A 61-million-person experiment in social influence and political mobilization. *Nature* 489(7415):295–298.
38. Ansolabehere S, Hersh E (2010) *The Quality of Voter Registration Records: A State-by-State Analysis. Report, Caltech/MIT Voting Technology Project* (California Institute of Technology, Pasadena, CA; Massachusetts Institute of Technology, Cambridge, MA).
39. Milkman KL, Beshears J, Choi JJ, Laibson D, Madrian BC (2011) Using implementation intentions prompts to enhance influenza vaccination rates. *Proc Natl Acad Sci USA* 108(26):10415–10420.

# Supporting Information

## Rogers et al. 10.1073/pnas.1524899113

*Best Letter (2008 voters)*

AMERICA VOTES

[Date]

Dear [Name],

You're a voter. **Thank you for voting in 2008!** Public records show that you voted in the most recent election. Again, thank you. It is **people like you** who decide the future of our country by getting out to the polls on **Election Day, November 2, 2010.**

This year, voters like you will join HUNDREDS of THOUSANDS of other [STATE] and cast a ballot on Election Day.

This election is important, and people like you will choose who represents us in Washington. You were a voter in 2008, you are still a voter.

**Voting takes a plan.**

☐ Will you vote on Election Day Tuesday November 2nd?_____

☐ What time will you vote? _____

☐ How you will get to your polling place? _____

**We hope the public record shows that you were a voter again** this election so we can thank you again.

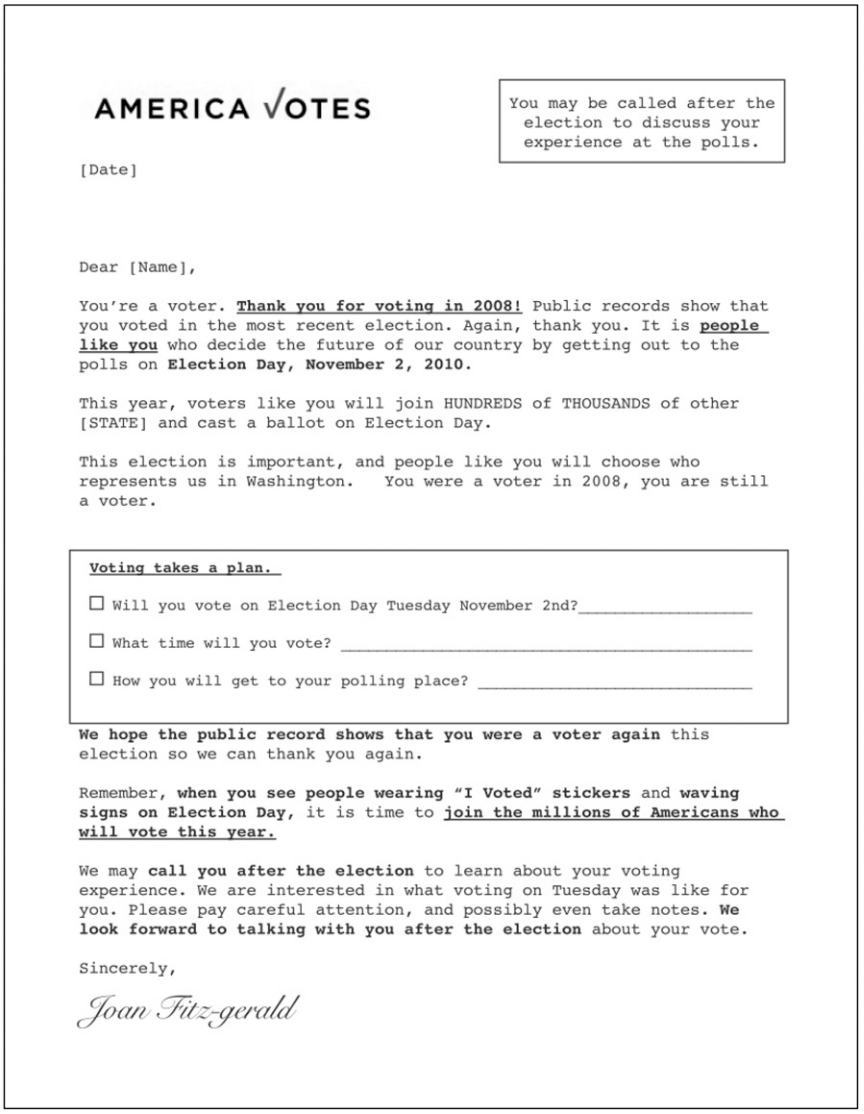
Remember, **when you see people wearing "I Voted" stickers** and **waving signs on Election Day,** it is time to **join the millions of Americans who will vote this year.**

Sincerely,

*Joan Fitz-gerald*

**Fig. S1.** (Continued)

*Best+Observable Letter (2008 voters)*



AMERICA VOTES

You may be called after the election to discuss your experience at the polls.

[Date]

Dear [Name],

You're a voter. **Thank you for voting in 2008!** Public records show that you voted in the most recent election. Again, thank you. It is **people like you** who decide the future of our country by getting out to the polls on **Election Day, November 2, 2010.**

This year, voters like you will join HUNDREDS of THOUSANDS of other [STATE] and cast a ballot on Election Day.

This election is important, and people like you will choose who represents us in Washington. You were a voter in 2008, you are still a voter.

**Voting takes a plan.**

☐ Will you vote on Election Day Tuesday November 2nd?_____

☐ What time will you vote? _____

☐ How you will get to your polling place? _____

**We hope the public record shows that you were a voter again** this election so we can thank you again.

Remember, **when you see people wearing "I Voted" stickers** and **waving signs on Election Day,** it is time to **join the millions of Americans who will vote this year.**

We may **call you after the election** to learn about your voting experience. We are interested in what voting on Tuesday was like for you. Please pay careful attention, and possibly even take notes. **We look forward to talking with you after the election** about your vote.

Sincerely,

*Joan Fitz-gerald*

**Fig. S1.** (Continued)

*Best Letter (2010 registrants)*

AMERICA VOTES

[Date]

Dear [Name],

**Thank you for registering to vote!** Public records show that you are newly registered for the upcoming election. Again, thank you. It is **people like you** who decide the future of our country by getting out to the polls on **Election Day, November 2, 2010.**

This year, voters like you will **join HUNDREDS of THOUSANDS of other [STATE]** and cast a ballot on Election Day.

This election is important, and people like you will choose who represents us in Washington. You were a voter in 2008, **you are still a voter.**

**Voting takes a plan.**

☐ Will you vote on Election Day Tuesday November 2nd?_____

☐ What time will you vote? _____

☐ How you will get to your polling place? _____

**We hope the public record shows that you are a voter in** this election so we can thank you again.

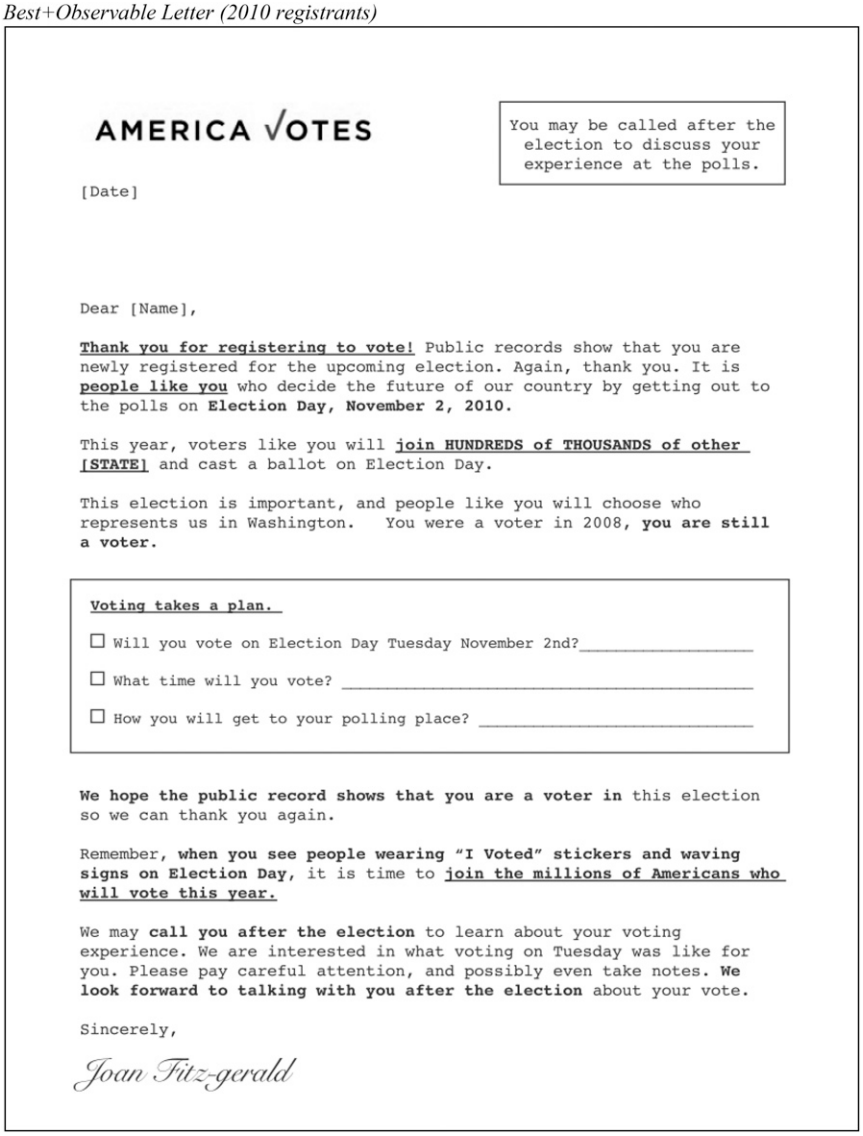
Remember, **when you see people wearing "I Voted" stickers and waving signs on Election Day,** it is time to **join the millions of Americans who will vote this year.**

Sincerely,

*Joan Fitz-gerald*

**Fig. S1.** (Continued)

*Best+Observable Letter (2010 registrants)*



AMERICA VOTES

You may be called after the election to discuss your experience at the polls.

[Date]

Dear [Name],

**Thank you for registering to vote!** Public records show that you are newly registered for the upcoming election. Again, thank you. It is **people like you** who decide the future of our country by getting out to the polls on **Election Day, November 2, 2010.**

This year, voters like you will **join HUNDREDS of THOUSANDS of other [STATE]** and cast a ballot on Election Day.

This election is important, and people like you will choose who represents us in Washington. You were a voter in 2008, **you are still a voter.**

**Voting takes a plan.**

☐ Will you vote on Election Day Tuesday November 2nd?_____

☐ What time will you vote? _____

☐ How you will get to your polling place? _____

**We hope the public record shows that you are a voter in** this election so we can thank you again.

Remember, **when you see people wearing "I Voted" stickers and waving signs on Election Day,** it is time to **join the millions of Americans who will vote this year.**

We may **call you after the election** to learn about your voting experience. We are interested in what voting on Tuesday was like for you. Please pay careful attention, and possibly even take notes. **We look forward to talking with you after the election** about your vote.

Sincerely,

*Joan Fitz-gerald*

**Fig. S1.** Full letters.

**Table S1. Balance check across demographics**

| Treatment group | % Female | % Democrat | % African American | 2008 primary turnout | 2006 general turnout | Average age* |
|---|---|---|---|---|---|---|
| Control | 59.78% | 55.14% | 15.28% | 32.04% | 80.66% | 43.34 |
| ($n$ = 76,963) | 46,008 | 42,437 | 11,760 | 24,659 | 62,078 | 76,498 |
| Nonobservable | 59.67% | 55.02% | 15.59% | 31.96% | 80.69% | 43.24 |
| ($n$ = 346,929) | 207,013 | 190,880 | 54,086 | 110,879 | 279,937 | 344,874 |
| Observable | 59.72% | 55.09% | 15.55% | 32.10% | 80.69% | 43.29 |
| ($n$ = 347,054) | 207,261 | 191,192 | 53,967 | 111,404 | 280,038 | 344,899 |

Multinomial logistic regression: LR chi2(12) = 8.96, $P$ = 0.7063.

*Age was missing for a small percentage of the observations.

**Table S2. Balance check across congressional districts**

| District | Control | Best | Best plus observable | Total |
|---|---|---|---|---|
| AZ 5 | 3,076 | 13,879 | 13,906 | 30,861 |
| | 3.82% | 3.83% | 3.83% | 3.83% |
| CA 11 | 2,801 | 12,459 | 12,485 | 27,745 |
| | 3.48% | 3.44% | 3.44% | 3.44% |
| FL 22 | 3,489 | 16,233 | 16,230 | 35,952 |
| | 4.34% | 4.48% | 4.47% | 4.46% |
| IA 3 | 3,325 | 14,378 | 14,354 | 32,057 |
| | 4.13% | 3.97% | 3.96% | 3.98% |
| IL 10 | 2,524 | 11,322 | 11,309 | 25,155 |
| | 3.14% | 3.12% | 3.12% | 3.12% |
| IL 14 | 2,374 | 10,808 | 10,807 | 23,989 |
| | 2.95% | 2.98% | 2.98% | 2.98% |
| IL 17 | 2,339 | 10,713 | 10,744 | 23,796 |
| | 2.91% | 2.95% | 2.96% | 2.95% |
| IN 9 | 3,204 | 14,767 | 14,762 | 32,733 |
| | 3.98% | 4.07% | 4.07% | 4.06% |
| MA 10 | 3,506 | 15,642 | 15,662 | 34,810 |
| | 4.36% | 4.31% | 4.32% | 4.32% |
| MD 1 | 3,221 | 14,424 | 14,427 | 32,072 |
| | 4.00% | 3.98% | 3.98% | 3.98% |
| MI 7 | 2,349 | 10,244 | 10,247 | 22,840 |
| | 2.92% | 2.83% | 2.83% | 2.83% |
| MI 9 | 2,841 | 12,955 | 12,965 | 28,761 |
| | 3.53% | 3.57% | 3.57% | 3.57% |
| MN 1 | 3,662 | 16,599 | 16,622 | 36,883 |
| | 4.55% | 4.58% | 4.58% | 4.58% |
| NC 8 | 3,418 | 15,713 | 15,717 | 34,848 |
| | 4.25% | 4.33% | 4.33% | 4.32% |
| ND 1 | 1,294 | 5,805 | 5,783 | 12,882 |
| | 1.61% | 1.60% | 1.59% | 1.60% |
| NY 19 | 2,877 | 12,982 | 12,978 | 28,837 |
| | 3.58% | 3.58% | 3.58% | 3.58% |
| NY 23 | 1,699 | 7,835 | 7,827 | 17,361 |
| | 2.11% | 2.16% | 2.16% | 2.15% |
| NY 24 | 2,888 | 13,361 | 13,374 | 29,623 |
| | 3.59% | 3.69% | 3.69% | 3.68% |
| OH 15 | 4,235 | 19,019 | 19,035 | 42,289 |
| | 5.26% | 5.25% | 5.25% | 5.25% |
| OH 18 | 3,485 | 15,735 | 15,770 | 34,990 |
| | 4.33% | 4.34% | 4.35% | 4.34% |
| PA 10 | 2,932 | 13,307 | 13,304 | 29,543 |
| | 3.64% | 3.67% | 3.67% | 3.67% |
| PA 11 | 1,782 | 8,123 | 8,113 | 18,018 |
| | 2.21% | 2.24% | 2.24% | 2.24% |
| PA 12 | 1,676 | 7,378 | 7,372 | 16,426 |
| | 2.08% | 2.03% | 2.03% | 2.04% |
| PA 15 | 2,258 | 9,797 | 9,802 | 21,857 |
| | 2.81% | 2.70% | 2.70% | 2.71% |
| PA 3 | 2,068 | 9,461 | 9,473 | 21,002 |
| | 2.57% | 2.61% | 2.61% | 2.61% |
| PA 7 | 2,941 | 12,926 | 12,930 | 28,797 |
| | 3.65% | 3.57% | 3.56% | 3.57% |
| SD 1 | 3,192 | 14,325 | 14,317 | 31,834 |
| | 3.97% | 3.95% | 3.95% | 3.95% |
| WI 7 | 2,761 | 12,474 | 12,476 | 27,711 |
| | 3.43% | 3.44% | 3.44% | 3.44% |
| WI 8 | 2,252 | 9,907 | 9,925 | 22,084 |
| | 2.80% | 2.73% | 2.74% | 2.74% |
| Did not vote in 2008 | 16,048 | 72,314 | 72,359 | 160,721 |
| | 19.94% | 19.94% | 19.95% | 19.95% |
| Voted in 2008 | 64,421 | 290,257 | 290,357 | 645,035 |
| | 80.06% | 80.06% | 80.05% | 80.05% |
| Total | 80,469 | 362,571 | 362,716 | 805,756 |

LR chi2(58) = 25.91; Pr = 1.000.

# Promoting Cooperation in the Field

Gordon Kraft-Todd[a], Erez Yoeli[b], Syon Bhanot[c], David Rand[d*]

February 18, 2015

**Abstract**

We review the growing literature of field experiments designed to promote cooperative behavior in policy-relevant settings outside the laboratory (e.g. conservation, charitable donations, voting). We focus on four categories of intervention that have been well studied. We find that material rewards and increased efficacy, interventions focused on altering the costs and benefits of giving, have at best mixed success. Social interventions based on observability and descriptive norms, conversely, are consistently highly effective. We then demonstrate how a theoretical framework based on reciprocity and reputation concerns explains why social interventions are typically more effective than cost-benefit interventions, and suggests ways to make cost-benefit interventions more effective. We conclude by discussing other less-studied types of intervention, and promising directions for future research.

[a]Yale University, Department of Psychology, Sheffield-Sterling-Strathcona Hall Room 206, 1 Prospect Street, New Haven CT 06511, gordon.kraft-todd@yale.edu
[b]Harvard University, Program for Evolutionary Dynamics, 1 Brattle Square Suite 6, Cambridge MA 02138, eyoeli@fas.harvard.edu
[c]Harvard University, John F. Kennedy School of Government, 79 John F. Kennedy Street, Cambridge MA 02138, syon_bhanot@hksphd.harvard.edu
[a]Yale University, Department of Psychology, Department of Economics, and School of Management, Sheffield-Sterling-Strathcona Hall Room 204, 1 Prospect Street, New Haven CT 06511, david.rand@yale.edu
*Corresponding author: (607) 592-0218

## Introduction

Many of society's biggest policy challenges—protecting the environment, providing healthcare, education, and safety, encouraging participation in the democratic process—are social dilemmas. These challenges require individuals to bear personal costs in order to benefit others, a behavior that is typically defined as "cooperation" [1]. There is a long tradition in both the social and natural sciences of studying cooperation theoretically using mathematical models and computer simulations, and of validating the theory empirically using laboratory experiments (for reviews, see Camerer and Fehr [2002] and Rand and Nowak [2013]). These lines of research are particularly exciting because, in addition to advancing scientific understanding, their results have the potential to provide insights into how to solve real-world social dilemmas.

It is often unclear, however, how to translate the findings of this (often abstract) literature to policy-relevant contexts. In particular, cost-effectiveness and practical feasibility are issues that are not typically relevant to (and thus not considered by) theory or lab experiments, but are essential for real-world applications. To bridge this gap, social science researchers have increasingly begun to perform *field* experiments exploring cooperation outside the laboratory. By using random assignment—the central tool of laboratory experiments—in the context of real-world social dilemmas, these studies enable researchers to draw clear conclusions about causality while also providing the external validity critical for policy recommendations.

Here, we provide an overview of this burgeoning literature investigating ways to promote real-world cooperation. We identify four categories of intervention that have been widely studied and summarize each (see Figure 1 for an overview). We then present a synthesis based on our theoretical understanding of the ultimate explanations for human cooperation. We show how this synthesis illuminates why some interventions usually succeed, and how it provides suggestions for ways to increase the effectiveness of others. Finally, we conclude with a discussion of other intervention categories which have been less thoroughly explored, and suggest directions for future work.

## Interventions to promote cooperation in the field

### *Cost-Benefit Interventions*

We begin by describing two classes of intervention rooted in a model of decision-making whereby people give because they derive some benefit from the outcomes of others (i.e. are "altruistic") [3]. From this theoretical perspective based on altruism, the choice of whether to cooperate involves weighing the cost to one's self against the benefit gained by others. Therefore these "Cost-Benefit Interventions" seek to change the (actual or perceived) costs and benefits of cooperation to increase its attractiveness: *material rewards* decrease the cost to the actor, and *increased efficacy* increases the benefits to the recipient.

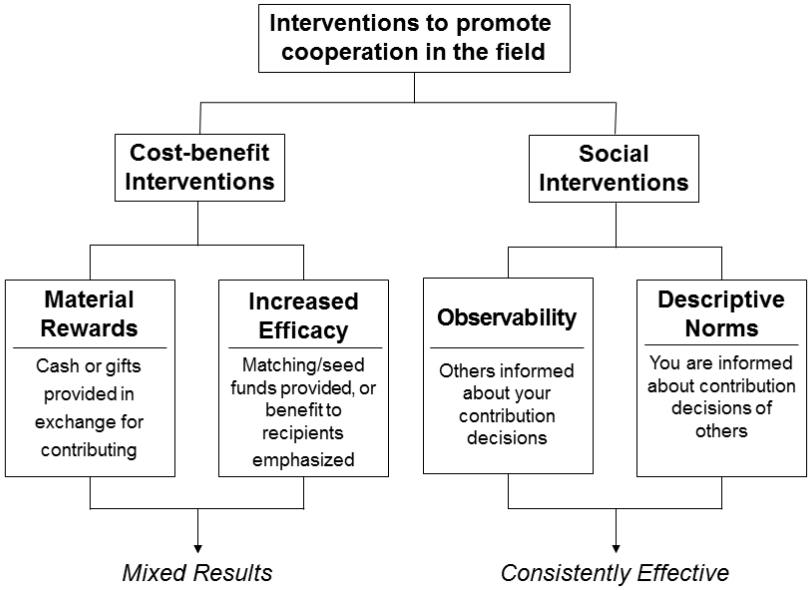



Figure 1. Summary of findings in this review

*Material rewards*

Some studies sought to decrease cooperation's cost to the self by offering material rewards in exchange for cooperating, such as cash, t-shirts or mugs, with mixed success [4-15]. For example, Landry et al [12] entered people who contributed to a fund-raiser into raffles to win a personal cash prize, and found a 47% increase in the amount of money raised relative to controls with no raffle. Lacetera, Macis and Slonim [2012], on the other hand, explored the effect of providing t-shirts at blood drives run by the Red Cross. They found that participation rates increased by about 25% at locations offering incentives, but that this increase was largely driven by participants that would have donated elsewhere instead traveling to locations that offered the reward. In the domain of energy, Yoeli, Hoffman, Rand and Nowak [2013] found that paying people $25 to sign up for a blackout prevention program had little effect on participation rates.

*Increased efficacy*

Other studies aimed to increase the perceived efficacy of contributing (i.e. the benefit created for others) either by supplementing donation amounts using matching or seed funds or by providing information that emphasizes the positive effects of contributing. Such efforts have also met with mixed success [12, 13, 16-24]. For example, Karlan and List [2007] found that offering matching funds increased donations to a charity by 19% compared to no matching funds, but that givers were insensitive to the size of the offered match; but Karlan, List and Shafir [2011] did not find any significant effect of matching grants on average giving. Seed money can increase perceived efficacy of giving by making it more likely that a fundraising goal is reached (a threshold effect), or by signaling an organization's quality (e.g. that the organization has been vetted by large, experienced donors). For example, List [2002] solicited charitable donations to purchase a $3,000 computer for a non-profit organization, and found that the average donation was more

than 7 times larger when potential donors were told that seed money had already paid for 67% of the goal, compared to 10% of the goal. However, Chen, Li and MacKie-Mason [2006] found that in a fundraising campaign for the Internet Public Library, advertising a seed donation that covered half of the fundraising goal ($10,000 out of $20,000) did not significantly increase contributions compared to a standard voluntary contribution mechanism (simply announcing a goal of $20,000). (Note that seed money may also have some element of descriptive norm information; see below.)

Thus, overall, the results are mixed. While Cost-Benefit Interventions may sometimes increase cooperation, they have been found to be ineffective in a number of cases. Therefore, it would be useful for policymakers to have other forms of intervention at their disposal.

## *Social Interventions*

We now turn to two classes of intervention which rely on social factors rather than material factors: *observability* makes the actor's behavior observable to others and *descriptive norms* provide information about others' behavior to the actor. While the simple economic model of decision-making described above (where people weigh material costs and benefits) would not predict these interventions to be effective, theoretical and experimental work from biology, psychology and behavioral economics suggests that they have promise (for a review, see Rand, Yoeli and Hoffman [2014]). An additional attractive feature of these "Social Interventions," relative to most Cost-Benefit Interventions, is that they are typically very inexpensive and easy to implement.

### *Observability*

Making one's contribution decision observable by others has consistently been found to increase cooperation [4, 6, 27-40]. For example, Yoeli, Hoffman, Rand and Nowak [2013] found that subjects were three times more likely to participate in a blackout prevention program when they enrolled by writing their names and apartment numbers on a publicly posted signup sheet, rather than just an anonymous ID number. Even subtle manipulations that only give the *impression* of being observed can increase cooperation. For example, posters of eyes have been found to increase honor-system payments for coffee in a university office by 276% (compared to images of flowers) [38], reduce the amount of litter left on university dining hall tables by 69% (compared to posters of flowers) [32], and increase money donated to charity collection buckets in a supermarket by 48% (compared to images of stars) [40].

### *Descriptive norms*

People are more likely to cooperate when they are told that others have cooperated, implying that cooperation is the social norm [10, 20, 23, 24, 35, 39, 41-58]. (Note that this type of intervention is the converse of observability: here you are informed about the behavior of others, rather than others being informed about your behavior.) For example, Frey and Meier [2004] increased the number of students contributing to a campus charity by 2.3% by informing them that 64% of students had contributed in the past (compared to informing them that 46% of students had contributed in the past). Goldstein, Cialdini and Griskevicius [2008] increased towel reuse by 9% in hotels by informing guests that 75% of previous

guests had reused their towels, compared to a standard environmental appeal (i.e. "Help Save the Environment"). This approach has been successfully applied in the energy domain by companies such as OPower and Enertiv, improving conservation by comparing customers' consumption to that of their peers (e.g., Allcott [2011], Ayres, Raseman and Shih [2012]). However, descriptive norms can also have perverse effects for some people: Bhanot [in prep] found that ranking consumers' water use relative to their neighbors may decrease conservation among those who conserved more than the norm. There is some evidence that this "backsliding" to the norm (known as the "boomerang effect") may be prevented by framing the rank ordering as a competition [53], or by messages about cooperating being the appropriate behavior (i.e. injunctive norms, as in Schultz, Nolan, Cialdini, Goldstein and Griskevicius [2007]).

**Synthesis: Reciprocity shapes human cooperation**

What explains why Social Interventions seem to be more effective than Cost-Benefit Interventions? And to what extent will the results of these specific field experiments generalize to other field settings? A theoretical understanding of human cooperation helps to answer these applied questions (and the patterns observed in these applications help to validate and extend our theoretical understanding) [26]. While there are many explanations for why people cooperate, we argue that the concept of reciprocity is particularly useful for organizing the literature on promoting cooperation in the field.

A key feature of human behavior is that future consequences often exist for your choices today. When interactions are repeated or reputations are at stake, cooperation can be in your long-run self-interest: it is worth paying the cost of cooperating today in order to earn the benefits of others' reciprocal cooperation with you in the future [1, 59, 60]. As a result, our preferences are shaped by reciprocity, and we typically develop reciprocally cooperative intuitions or "social heuristics" [1, 61-64]. Thus, although people may not always explicitly *deliberate* over the impact of their actions on their reputations, reciprocal concerns are deeply rooted in human psychology and influence our intuitive, gut responses.

This theoretical account of human prosociality makes predictions regarding which interventions will work better than others: those interventions that best engage people's sense of reciprocity should be most effective. Indeed, the field experiments reviewed here fit this pattern. The highly effective Social Interventions strongly invoke reciprocity. *Observability* engages subjects' reputational concerns by allowing others to better observe—and thus reciprocate—their good deeds. And *Descriptive norms* engage reciprocal concerns by providing information about how others have acted, and therefore what others are likely to expect of you (i.e. which of your actions will be rewarded and punished).

Conversely, the Cost-Benefit Interventions that met with only mixed success do not engage reciprocity and reputation, or even worse, sometimes undercut these concerns. *Material rewards* for being cooperative can "crowd out" the reputational benefits that typically come with contributing [7, 65, 66]: they make it unclear whether contributions were made because you are actually a cooperative person (and thus deserving of a good reputation, both in the eyes of others and of yourself), or just for the selfish purpose of receiving the material reward [67]. This perspective suggests that material rewards that benefit other people as well as the cooperator might be more effective, because they may seem less indicative of

a self-interested motive; for example, a party for the team that raises the most money in a fundraiser, or the suite that uses the least electricity in a dorm. *Increased efficacy* has two issues from a reciprocity perspective. First, the cost of one's cooperation is typically much easier for others to observe than the beneficial effects, as those benefits typically occur later, and are more diffuse and are harder to quantify. Second, increased efficacy of your contribution arising from donation matching may not feel attributable to you, but instead to those who contributed the match money. For both of these reasons, increased efficacy may not bring greater reputational gains (or lead to one feeling like a better person for having contributed). This perspective offers a potential solution: make efficacy of contributions publicly observable to others. For example, when listing the amount people donated to a cause, include the match amount in each individual's donation total.

This reciprocity framework also sheds light on whether, and when, these interventions will be effective in contexts beyond those in which they were tested. Reciprocity and reputation are dominant features of human social interaction across settings [68]. Thus we expect interventions based on these principles to be widely effective. This is particularly true in settings where reputational concerns are greatest, that is, when we have particularly valuable relationships with those who can observe our behavior. For example, the blackout reduction study of Yoeli, Hoffman, Rand and Nowak [2013] found that observability had a much bigger effect among apartment owners (who typically have long-term relationships with their neighbors) than among renters (who are more transient).

Further implications arise from the fact that reciprocity and reputation concerns may often be operating at an intuitive, rather than explicitly conscious, level [1, 61-64, 69, 70]: interventions that more heavily engage intuitive, emotional processes may be more effective in promoting cooperation. Consistent with this prediction, Small, Loewenstein and Slovic [2007] found that people were more willing to donate to emotional salient "identifiable victims" than to causes described with rationally compelling statistics. Furthermore, subjects in their experiments donated more to identifiable victims when primed to make their decision emotionally or "go with their gut," and subjects in the economic cooperation games experiments of Rand, Greene and Nowak [2012] and Rand, Peysakhovich, Kraft-Todd, Newman, Wurzbacher, Nowak and Greene [2014] contributed more to the public good when forced to decide more intuitively. Such results provide another reason Cost-Benefit Interventions may sometimes be ineffective: these interventions typically aim to change the results of conscious, deliberative calculations regarding costs and benefits, rather than appealing to donors' intuitions.

Our theoretical framework also suggests important limitations to the Social Interventions. One must beware not to "crowd out" cooperation by making reputational rewards too explicit: in the same way that material rewards can suggest selfish motives for cooperation, so too can explicit reputational rewards [69]. Additionally, reciprocity and reputation concerns will only motivate cooperation if cooperating is typically perceived as desirable: in communities which disparage cooperation in a particular domain, many of these interventions are unlikely to work (e.g. fund raising for the National Rifle Association in politically liberal communities, or for environmental sustainability in politically conservative communities) [27, 50, 73].

## Future directions

In addition to the four categories of intervention we have discussed here, numerous other approaches to promoting cooperation have been explored in the field. These include non-contingent gifts to induce reciprocal feelings of obligation [22, 39, 74-78]; setting defaults such that non-cooperation requires actively opting out [16, 42, 79]; solicitations explicitly asking people to cooperate [15, 80-83]; the framing of such solicitations [84]; variation of the characteristics of the people making such solicitations [13, 35, 37]; participatory decision-making, whereby cooperators get to give input on what public goods are produced [36, 37]; and instrumental information enabling cooperation (e.g. real-time feedback on home energy use) [10, 12, 48-51, 57, 58, 85]. Expanding the policymaker's toolkit via further exploration of these and other potential interventions is a critical direction for future research on human cooperation. In doing so, the theoretical perspective we present here can help to illuminate which approaches are particularly promising, and provide guidance on how to optimize their effectiveness.

Finally, we end by suggesting one additional avenue for further investigation. A topic that has received little attention in the context of field experiments on cooperation is the formation and modification of habits. Rather than one-time actions, the solutions to many real-world public goods require long-term behavior modification [86]. A large of body of evidence from social and cognitive psychology suggests that we internalize behaviors that are typically successful, and adopt them as intuitive default responses [87-89]. Thus particularly successful interventions will help to overcome habitual inertia, further increasing the initial gains. Understanding which interventions most effectively build cooperative habits, and what factors contribute to treatment persistence more generally, is of great importance for effecting real-world change.

**Conflict of interest statement**
Nothing declared.

**Acknowledgements**
We thank Lucia Chen, Ruchita Gupta, Grant Kopplin, Laura Peng, and especially Vivien Caetano for their invaluable research assistance, and gratefully acknowledge financial support from the John Templeton Foundation.

## References

1. D. G. Rand, M. A. Nowak, Human Cooperation. *Trends in Cognitive Sciences* **17**, 413 (2013).

2. C. F. Camerer, E. Fehr, Measuring social norms and preferences using experimental games: A guide for social scientists. (2002).

3. G. S. Becker, Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology. *Journal of Economic Literature* **14**, 817 (1976).

4. N. Ashraf, O. Bandiera, K. Jack, *No Margin, No Mission?: A Field Experiment on Incentives for Pro-social Tasks*. (Centre for Economic Policy Research, 2012).
5. C. Noussair, D. van Soest, J. Stoop, Punishment, reward, and cooperation in a framed field experiment. (2011).

6. E. Yoeli, M. Hoffman, D. G. Rand, M. A. Nowak, Powering up with indirect reciprocity in a large-scale field experiment. *Proceedings of the National Academy of Sciences* **110**, 10424 (2013).

**The authors explore the role of observability in increasing sign-ups for a voluntary demand response program (where energy companies are able to moderate power availability during peak use times). In an intervention of 2400 subjects, they find that including particpants' names on the sign-up sheet (in addition to a subject ID number) increases sign-ups threefold. This is especially for those whose living arrangements imply greater reputational concerns, i.e. owners (rather than renters) and residents of apartment buildings (rather than row houses or individual homes).

7. U. Gneezy, A. Rustichini, Pay Enough or Don't Pay at All. *Quarterly Journal of Economics* **115**, 791 (2000).

8. B. Uri Gneezy, Aldo Rustichini, A Fine Is a Price. *The Journal of Legal Studies* **29**, 1 (2000).

9. L. Goette, A. Stutzer, "Blood donations and incentives: Evidence from a field experiment" (IZA Discussion Papers, 2008).
10. P. Dolan, R. Metcalfe, Neighbors, knowledge, and nuggets: two natural field experiments on the role of incentives on energy conservation. (2013).

11. N. Lacetera, M. Macis, R. Slonim, Will There Be Blood? Incentives and Displacement Effects in Pro-social Behavior. *American Economic Journal: Economic Policy* **4**, 186 (2012).

**This paper reports on a field experiment of 72 blood drives (and aggregate data from close to 14,000) showing that material incentives (e.g. t-shirts, mugs) increase the number of blood donors at a drive by about 25%. They also find, however, that drives that have incentives displace participation in nearby and proximately-timed drives. They conclude that while material incentives increase blood donations, the effect may often be over-estimated.

12. C. Landry, A. Lange, J. A. List, M. K. Price, N. G. Rupp, "Toward an understanding of the economics of charity: Evidence from a field experiment" (National Bureau of Economic Research, 2005).
13. J. A. List, M. K. Price, The role of social connections in charitable fundraising: Evidence from a natural field experiment. *Journal of Economic Behavior & Organization* **69**, 160 (2009).

14. E. Xiao, D. Houser, "Sign Me Up! A Model and Field Experiment on Volunteering" (2014).
15. K. I. Ito, Takanori; Tanaka, Makoto, The Persistence of Intrinsic and Extrinsic Motivation: Experimental Evidence from Energy Demand. *Available at: http://www.abef.jp/event/2014/pdf_abst/PR0005.pdf*, (working paper).

16. D. Karlan, J. A. List, Does Price Matter in Charitable Giving? Evidence from a Large-Scale Natural Field Experiment. *American Economic Review* **97**, 1774 (2007).

**Using over 50,000 subjects in a direct-mail fundrasing campaign for a non-profit, the authors test the effect of different matching grant ratios ($1:$1, $2:$1, $3:$1) on charitable giving. They find that while a matching grant does increase donations (about 15% per solicitation), the size of the match does not signficantly increase giving.

17. D. Rondeau, J. List, Matching and challenge gifts to charity: evidence from laboratory and natural field experiments. *Exp Econ* **11**, 253 (2008).

18. S. Meier, Do subsidies increase charitable giving in the long run? Matching donations in a field experiment. *Journal of the European Economic Association* **5**, 1203 (2007).

19. U. Gneezy, E. A. Keenan, A. Gneezy, Avoiding overhead aversion in charity. *Science* **346**, 632 (2014).

**This paper tests a novel intervention in charitable fundraising, examining the effect of informing prospective donors that their money will go directly to charitable programs rather than overhead costs. They find that using this intervention in a direct-mail campaign to 40,000 potential donors raises 75% more money than a standard seed donation treatment and 89% more than a match treatment.

20. C. E. Cryder, G. Loewenstein, H. Seltman, Goal gradient in helping behavior. *Journal of Experimental Social Psychology* **49**, 1078 (2013).

21. D. Karlan, J. A. List, E. Shafir, Small matches and charitable giving: Evidence from a natural field experiment. *Journal of Public Economics* **95**, 344 (2011).

22. Y. Chen, X. Li, J. K. MacKie-Mason, Online fund-raising mechanisms: A field experiment. *Contributions in Economic Analysis & Policy* **5**, (2006).

23. R. a. M. Koning, Jacob, Social Influence Mechanisms in Crowdfunding Cascades: When Nothing is Better than Something. *Available at SSRN: http://ssrn.com/abstract=2308161*, (2013).

24. R. Martin, J. Randal, How is donation behaviour affected by the donations of others? *Journal of Economic Behavior & Organization* **67**, 228 (2008).

25. J. List, Lucking-Reiley, David, The Effects of Seed Money and Refunds on Charitable Giving: Experimental Evidence from a University Capital Campaign. *Journal of Political Economy* **110**, 215 (2002).

26. D. G. Rand, E. Yoeli, M. Hoffman, Harnessing Reciprocity to Promote Cooperation and the Provisioning of Public Goods. *Policy Insights from the Behavioral and Brain Sciences* **1**, 263 (2014).

***This paper reviews studies in which manipulations of reciprocity are used to increase cooperation. The authors explain the theoretical background of these interventions in game theory and review five interventions relevant to policy-makers: maknig behavior observable, focusing on effectiveness, not letting people 'avoid the ask', projecting a norm of cooperation, and creating habits of cooperation.

27. D. Ariely, A. Bracha, S. Meier, Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially. *American Economic Review* **99**, 544 (2009).

28. M. Ekström, Do watching eyes affect charitable giving? Evidence from a field experiment. *Exp Econ* **15**, 530 (2012).

29. A. R. Soetevent, Anonymity in giving in a natural context—a field experiment in 30 churches. *Journal of Public Economics* **89**, 2301 (2005).

30. N. Lacetera, M. Macis, Social image concerns and prosocial behavior: Field evidence from a nonlinear incentive scheme. *Journal of Economic Behavior & Organization* **76**, 225 (2010).

31. D. Karlan, M. A. McConnell, "Hey look at me: The effect of giving circles on giving" (National Bureau of Economic Research, 2012).

32. M. Ernest-Jones, D. Nettle, M. Bateson, Effects of eye images on everyday cooperative behavior: a field experiment. *Evolution and Human Behavior* **32**, 172 (2011).

33. D. Francey, R. Bergmüller, Images of Eyes Enhance Investments in a Real-Life Public Good. *PLoS ONE* **7**, e37397 (2012).

34. P. Funk, Social incentives and voter turnout: evidence from the swiss mail ballot system. *Journal of the European Economic Association* **8**, 1077 (2010).

35. G. D'Adda, "Leadership and influence: Evidence from an artefactual field experiment on local public good provision" (Working Paper Series, Department of Economics, University of Zurich, 2011).

36. B. A. Olken, Direct Democracy and Local Public Goods: Evidence from a Field Experiment in Indonesia. *The American Political Science Review* **104**, 243 (2010).

37. C. Cavalcanti, F. Schläpfer, B. Schmid, Public participation and willingness to cooperate in common-pool resource management: A field experiment with fishing communities in Brazil. *Ecological Economics* **69**, 613 (2010).

38. M. Bateson, D. Nettle, G. Roberts, *Cues of being watched enhance cooperation in a real-world setting*. (2006), vol. 2, pp. 412-414.

39. F. Alpizar, F. Carlsson, O. Johansson-Stenman, Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica. *Journal of Public Economics* **92**, 1047 (2008).

40. K. L. Powell, G. Roberts, D. Nettle, L. Fusani, Eye Images Increase Charitable Donations: Evidence From an Opportunistic Field Experiment in a Supermarket. *Ethology* **118**, 1096 (2012).

41. R. Martin, J. Randal, Voluntary contributions to a public good: a natural field experiment. *Unpublished manuscript, Victoria University, New Zealand*, (2005).

42. F. Carlsson, O. Johansson-Stenman, N. Pham Khanh, Funding a new bridge in rural Vietnam: A field experiment on conditional cooperation and default contributions. (2011).

43. R. Croson, J. Shang, The impact of downward social information on contribution decisions. *Exp Econ* **11**, 221 (2008).

44. R. Croson, F. Handy, J. Shang, Keeping up with the Joneses: The relationship of perceived descriptive social norms, social information, and charitable giving. *Nonprofit Management and Leadership* **19**, 467 (2009).

45. J. Shang, R. Croson, A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods. *Economic Journal* **119**, 1422 (2009).

46. B. S. Frey, S. Meier, Social Comparisons and Pro-Social Behavior: Testing "Conditional Cooperation" in a Field Experiment. *The American Economic Review* **94**, 1717 (2004).

47. N. J. Goldstein, R. B. Cialdini, V. Griskevicius, A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels. *Journal of Consumer Research* **35**, 472 (2008).

**The authors test the effect of descriptive norm information on hotel towel reuse in 190 rooms over 80 days. They find that a door hanger that asks guests to reuse their towels which includes information about others' compliance (e.g. "Almost 75% of guests who are asked to participate in our new resource savings program do help by using their towels more than once.") increases towel reuse by 9% over a generic message to "Help save the environment."

48. H. Allcott, Social norms and energy conservation. *Journal of Public Economics* **95**, 1082 (2011).

**This paper reports the result of home energy score mailers on reducing home energy use in over 600,000 homes. The home energy report includes information comparing the resident's energy use to their neighbors and tips to conserve energy. This experiment resulted in a small but robust decrease in energy use for homes receiving home energy scores.

49. I. Ayres, S. Raseman, A. Shih, Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage. *Journal of Law, Economics, and Organization*, (2012).

50. D. L. Costa, M. E. Kahn, ENERGY CONSERVATION "NUDGES" AND ENVIRONMENTALIST IDEOLOGY: EVIDENCE FROM A RANDOMIZED RESIDENTIAL ELECTRICITY FIELD EXPERIMENT. *Journal of the European Economic Association* **11**, 680 (2013).

51. D. A. Brent, J. H. Cook, S. Olsen, Norms and Water Conservation: How Do Effects Vary Across and Within Utilities? , (2014).

52. A. S. Gerber, T. Rogers, Descriptive Social Norms and Motivation to Vote: Everybody's Voting and so Should You. *The Journal of Politics* **71**, 178 (2009).

53. S. Bhanot, Rank and Response: A Field Experiment on Peer Information and Water Use Behavior. *Available at: http://goo.gl/AlWbex*, (in prep).

**This paper reports the result of an experiment of mailers comparing residents' water use to their neighbors in over 5000 households. Though cooperative and competitively-framed messages do not demonstrate treament effects different than control messages, the competitive conditions resulted in interesting heterogeneities: high users reduced their water use while low users increased their water use.

54. P. W. Schultz, J. M. Nolan, R. B. Cialdini, N. J. Goldstein, V. Griskevicius, The Constructive, Destructive, and Reconstructive Power of Social Norms. *Psychological Science (Wiley-Blackwell)* **18**, 429 (2007).

55. E. L. Paluck, Reducing intergroup prejudice and conflict using the media: a field experiment in Rwanda. *Journal of personality and social psychology* **96**, 574 (2009).

56. E. L. Paluck, H. Shepherd, The salience of social referents: A field experiment on collective norms and harassment behavior in a school social network. *Journal of personality and social psychology* **103**, 899 (2012).

57. P. J. Ferraro, M. K. Price, Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment. *Review of Economics and Statistics* **95**, 64 (2013).

58. H. Allcott, T. Rogers, The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation. *American Economic Review* **104**, 3003 (2014).

59. M. A. Nowak, K. Sigmund, Evolution of indirect reciprocity. *Nature* **437**, 1291 (2005).

60. D. Fudenberg, E. S. Maskin, The Folk Theorem in Repeated Games with Discounting or with Incomplete Information. *Econometrica* **54**, 533 (1986).

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

62. D. G. Rand *et al.*, Social Heuristics Shape Intuitive Cooperation. *Nature Communications* **5**, 3677 (2014).

63. V. Grimm, F. Mengel, Let me sleep on it: Delay reduces rejection rates in ultimatum games. *Economics Letters* **111**, 113 (2011).

64. M. Sutter, M. Kocher, S. Straub, Bargaining under time pressure in an experimental ultimatum game. *Economics Letters* **81**, 341 (2003).

65. B. S. Frey, F. Oberholzer-Gee, The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out. *American Economic Review* **87**, 746 (1997).

66. G. E. Newman, D. M. Cain, Tainted Altruism: When Doing Some Good Is Evaluated as Worse Than Doing No Good at All. *Psychological Science* **25**, 648 (2014).

67. R. BÉNabou, J. Tirole, Incentives and Prosocial Behavior. *American Economic Review* **96**, 1652 (2006).

68. E. Ostrom, Toward a behavioral theory linking trust, reciprocity, and reputation. (2003).

69. M. H. Hoffman, Christian; Nowak, Martin, Modesty. (Working Paper).

70. M. Hoffman, E. Yoeli, M. A. Nowak, Cooperate without looking: Why we care what people think and not just what they do. *Proceedings of the National Academy of Sciences* **112**, 1727 (2015).

71. D. A. Small, G. Loewenstein, P. Slovic, Sympathy and callousness: The impact of deliberative thought on donations to identifiable and statistical victims. *Organizational Behavior and Human Decision Processes* **102**, 143 (2007).

72. D. G. Rand *et al.*, Social Heuristics Shape Intuitive Cooperation. *Nature Communications* **5**, (2014).

73. T. Ellingsen, B. Herrmann, M. A. Nowak, D. G. Rand, C. E. Tarnita, "Civic capital in two cultures: the nature of cooperation in Romania and USA" (CESifo Working Paper: Behavioural Economics, 2012).

74. A. J. Caplan, D. Jackson-Smith, S. Marquart-Pyatt, Does 'free-sampling' enhance the value of public goods? *Applied Economics Letters* **17**, 335 (2010).

75. T. de Hoop, L. van Kempen, R. Fort, Do Cash Transfers Crowd Out Community Investment in Public Goods? Lessons from a Field Experiment on Health Education. *Nonprofit and Voluntary Sector Quarterly* **41**, 232 (2012).

76. A. Falk, Gift Exchange in the Field. *Econometrica* **75**, 1501 (2007).

77. U. Gneezy, J. A. List, Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments. *Econometrica* **74**, 1365 (2006).

78. A. Gneezy, U. Gneezy, L. D. Nelson, A. Brown, Shared Social Responsibility: A Field Experiment in Pay-What-You-Want Pricing and Charitable Giving. *Science* **329**, 325 (2010).

79. D. Pichert, K. V. Katsikopoulos, Green defaults: Information presentation and pro-environmental behaviour. *Journal of Environmental Psychology* **28**, 63 (2008).

80. J. Andreoni, J. M. Rao, H. Trachtman, "Avoiding the ask: a field experiment on altruism, empathy, and charitable giving" (National Bureau of Economic Research, 2011).

81. M. Knutsson, P. Martinsson, C. Wollbrant, Do people avoid opportunities to donate?: A natural field experiment on recycling and charitable giving. *Journal of Economic Behavior & Organization* **93**, 71 (2013).

82. N. Malhotra, M. R. Michelson, T. Rogers, A. A. Valenzuela, Text Messages as Mobilization Tools: The Conditional Effect of Habitual Voting and Election Salience. *American Politics Research* **39**, 664 (2011).

83. D. W. Nickerson, T. Rogers, Do You Have a Voting Plan?: Implementation Intentions, Voter Turnout, and Organic Plan Making. *Psychological Science* **21**, 194 (2010).

84. C. J. Bryan, G. M. Walton, T. Rogers, C. S. Dweck, Motivating voter turnout by invoking the self. *Proceedings of the National Academy of Sciences* **108**, 12653 (2011).

85. S. Houde, A. Todd, A. Sudarshan, J. A. Flora, K. C. Armel, Real-time Feedback and Electricity Consumption. *Economics* **95**, 1082 (2013).

86. E. Frey, T. Rogers, Persistence How Treatment Effects Persist After Interventions Stop. *Policy Insights from the Behavioral and Brain Sciences* **1**, 172 (2014).

87. G. Gigerenzer, D. G. Goldstein, Reasoning the fast and frugal way: models of bounded rationality. *Psychological review* **103**, 650 (1996).

88. D. Kahneman, A perspective on judgment and choice: mapping bounded rationality. *American psychologist* **58**, 697 (2003).

89. A. Peysakhovich, D. G. Rand, Habits of virtue: creating norms of cooperation and defection in the laboratory. *Management Science*, (in press).