Vorlaufer, T. (2019). Effects of double-anonymity on pro-and anti-social behavior: Experimental evidence from a lab in the field. Journal of Behavioral and Experimental Economics, 81, 216-225. https://doi.org/10.1016/j.socec.2019.07.001

Effects of Double-Anonymity on Pro- and Anti-Social Behavior: Experimental Evidence from a Lab in the Field

Tobias Vorlaufer [a b]

[a] School of Business and Economics, Marburg University, Am Plan 2, 35032 Marburg, Germany

[b] Institute of Environmental Systems Research, Osnabrück University, Barbarastr. 12, 49076 Osnabrück, Germany

Accepted Version

Abstract

This paper examines whether different degrees of subject-experimenter anonymity influence pro- and anti-social behavior in lab-in-the-field experiments. To do this, a Dictator Game (DG) and a Joy-of-Destruction Mini-Game (JoD) were conducted with 480 subjects in rural Namibia. In addition to a strict double-anonymous treatment two single-anonymous treatments are introduced. One of them involves the disclosure of decisions directly to the experimenter. Thereby, it is possible to disentangle the effect of pure double-anonymity from the procedure of the decision-making. The presented results carry relevant implications for a methodologically sound implementation of lab-in-the-field experiments. Both in the DG and JoD, strict double-anonymous procedures do not produce significantly different behavior than under single-anonymity. Whether revealing decisions personally to the experimenter influences individual behavior cannot be consistently answered with the obtained results. The personal disclosure leads to significantly more pro-social and less anti-social behavior in one out of two treatment comparisons. From a conservative perspective, researchers are however advised to assure sufficient privacy for subjects from experimenters during the decision-making process.

Keywords: Field experiments, dictator game, joy-of-destruction, double anonymity, experimenter-demand effect

Funding: This research was funded by the Southern African Science Service Center for Climate Change and Adaptive Land-Use Management (SASSCAL) through the German Federal Ministry for Education and Research [grant number 01LG1201B].

Acknowledgments: The study would have been impossible without the dedicated work of all research assistants and the hospitality of the participating communities. Special thanks go to Loukas Balafoutas, Thomas Falk, Christian Hönow, Michael Kirk, Ivo Steimanis and Björn Vollan and to the participants of the "12th Annual Meeting of the Environment for Development Initiative" and "3rd Annual Workshop on Experimental Economics for the Environment" for their valuable feedback and discussions.
An earlier version of the manuscript was published as a chapter in the author's dissertation: Vorlaufer, T. (2018). Three Essays on Environment and Development: A Behavioral Perspective (Doctoral dissertation). Philipps-Universität Marburg, Germany.

1. Introduction

Economic experiments are increasingly conducted in field settings, many in developing countries. Several unique characteristics of lab-in-the-field experiments suggest that they are especially likely to be vulnerable to social Experimenter Demand Effects (EDE) (Zizzo 2010). The social distance between researchers and subjects is commonly larger; often accompanied by larger perceived status differences between subjects and experimenters. Cilliers et al. (2015) found for example that the presence of white foreigners significantly increases pro-social behavior in a field setting in Sierra Leone. Moreover, lab-in-the-field experiments cannot rely on permanent infrastructure to recruit and run experiments. As a result, experimenters have commonly more face-to-face interactions with subjects, and often cooperate with organizations to recruit subjects[1]. One option to reduce social EDE are procedures that assure experimenter-subject anonymity throughout the experiment. This paper contributes to the methodological foundations of lab-in-the-field experiments by evaluating to what extent varying degrees of experimenter-subject anonymity affect social EDE. To do this, a Dictator Game (DG) and a Joy of Destruction Mini-Game (JoD) were conducted in rural Namibia. The results indicate that ID numbers, linking individual decisions to socio-economic survey data, alone, do not induce *additional* social EDE compared to full double-anonymity. However, revealing decisions personally to the experimenter decreases destructive behavior and increases transfer in the DG.

Subjects' adaptation of their behavior (consciously or unconsciously), due to the awareness that they are observed, is generally known as the "Hawthorne Effect" or the observer effect (Levitt and List 2011). Also known in the experimental literature as EDE, the effect describes a behavioral change of subjects "due to cues about what constitutes appropriate behavior (behavior 'demanded' from them)" (Zizzo 2010, 75). EDE are a particular concern, if positively correlated with the true experimental objective (Zizzo 2010). On the one hand, certain treatments may induce stronger EDE compared to a control treatment, thus biasing the estimation of the treatment effect itself. One example are experiments that vary group compositions to measure discriminative behavior (e.g. Chakravarty et al. 2016). On the other hand, certain sub-samples may be more susceptible for social EDE. For example, individuals affected by natural disasters may want to behave more "appropriately" motivated by attracting support (cf Cilliers,

[1] This constitutes one major difference to online experiments, which lack direct interactions between subjects and experimenters. De Quidt et al. (2018) show that EDE are modest in terms of their effect sizes for such experiments.

Dube, and Siddiqi 2015). This can in turn bias effect estimates of such events on social preferences (e.g. Cassar, Healy, and von Kessler 2017).

The concept of social EDE relates closely to the social desirability bias that has been studied comprehensively in the context of psychology and social sciences. Besides behavioral data from experiments, this research also focuses on survey responses[2]. Multiple methods have been proposed to either control for or reduce this bias (Nederhof 1985). For example, social desirability scales measure through survey items individuals' susceptibility for social desirable responses (Crowne and Marlowe 1960). The random response techniques is one method to directly minimize the social desirability bias by concealing individual answers from the researcher (Warner 1965). One key difference is that social-desirability biases can originate from conscious or unconscious other-deception and/or from unconscious self-deception (Millham and Kellogg 1980). In the latter form, respondents actually believe in their given responses. Consequently, methods that conceal individual responses from the researcher are ineffective in reducing any social desirability bias stemming from self-deception.

In the context of experimental research, assuring experimenter-subject anonymity is one among many mechanisms used particularly to reduce social EDE[3]. However, several characteristics of lab-in-the-field experiments render the implementation of subject-experimenter anonymity more difficult than in a controlled lab environment, and are likely to create additional costs. Lab-in-the-field experiments are commonly conducted in venues that do not provide the same level of privacy as labs. In addition, non-standard subject pools are commonly less-educated and many subjects have - especially in developing countries - low literacy skills. Therefore, experimenters often directly assist, observe and/or record decisions. Due to the greater variance

[2] Klein et al. (2012) provide a historical overview on the discussion and practices with regard to "experimenter bias" and "demand characteristics" in experimental psychology. Rosenthal & Rosnow (2009) provide an extended account of psychological research on "experimenter effects".

[3] Other methods to minimize EDE include for example sufficiently large monetary stakes, a between-subject design and non-deceptive obfuscation (Zizzo 2010). It is important to recognize that all procedures that aim to reduce social EDE, cannot rule out that social EDE are induced. One may classify social EDE as strategic and non-strategic. In the former case, subjects' behavior may be driven by the anticipation of positive or negative consequences contingent on their behavior in the experiment. Double-anonymity potentially reduces such social EDE at the individual level, since decisions cannot be attributed to individuals. Social EDE may be driven, however, also by the desire of subjects to exhibit socially desirable behavior at an aggregated level (e.g. the group- or village-level). This driver of social EDE cannot be addressed by double-anonymity, if experimental sessions do not include subjects from different groups or villages. Likewise, double-anonymity cannot minimize non-strategic social EDE. Subjects drawing conclusions on the purpose and objective of the research might want (unconsciously or consciously) to meet the expectations of the researchers (for example reciprocate the opportunity to participate in the experiment).

of socio-economic characteristics in non-standard subject pools, it is usually desirable to connect experimental data to information from post-experiment questionnaires. This requirement adds additional complexity to double-anonymous procedures in the field; especially if questionnaires have to be administered by enumerators.

From a methodological perspective it is therefore especially important to better understand and systematically investigate whether, and to what extent, current experimental procedures and practices influence behavior in lab-in-the-field experiments. One key question is which procedures successfully minimize social EDE. A number of lab experiments have investigated effects of subject-experimenter anonymity without clearly finding support for or against it. In the following, I will refer to subject-subject anonymity[4] as single-anonymity and subject-experimenter with subject-subject anonymity as double- anonymity (cf Barmettler, Fehr, and Zehnder 2012). Some studies refer to these categories as single- and double-blind respectively. Most studies in this context have been conducted with the DG, as it allows implementation of double-anonymity protocols with relative ease. While two studies find that transfers in a single-anonymity condition are higher than under double-anonymity (Cardenas 2014; Hoffman et al. 1994), other studies find no evidence for anonymity effects (Barmettler, Fehr, and Zehnder 2012; Bolton, Katok, and Zwick 1998; Cadsby, Servátka, and Song 2010; Hoffman, McCabe, and Smith 1996) or only among certain samples (Lesorogol and Ensminger 2014).

To my knowledge only two studies compare double-anonymity and single-anonymity procedures in the field. Cardenas (2014) conducted DG in rural Colombia, where subjects under single-anonymity revealed their decision directly to the experimenter. While he finds significant lower levels of generosity under double-anonymity, this effect may be confounded by cross-talk, since the double-anonymity sessions were conducted after the single-anonymity sessions in the same village. Lesorogol & Ensminger (2014) conducted the DG as lab-in-the-field experiment in one community in the US and two communities in Kenya. Subjects in their US sample reduce pro-social behavior under double-anonymity, which is however not observed in Kenya. The lack of detailed information regarding the single-anonymity procedures does not allow the reader to infer how decisions were made in this treatment (in private or with the experimenter present).

[4] In economic experiments subject-subject anonymity is usually assured due to ethical concerns. Especially in the field, where subjects of the same session are likely to interact afterwards, lifting subject-subject anonymity may induce serious conflicts. In specific cases, experimental designs may require to indicate with whom subjects are matched in the experiment. In the case of 2-player experiments, a random mechanism can be implemented that masks the individual decisions and maintains subject-subject anonymity (e.g. the "hidden" treatment in Abbink & Herrmann (2011)).

This paper's contribution to the existing research on double-anonymity is threefold and thereby tries to go beyond the two studies presented above. First, a strict double-anonymity treatment – DOUBLE - is compared with one single-anonymity procedure, which is designed to be as similar to the double-anonymity procedure as possible - referred to as IDNUM. The experimental procedures here were exactly as in the DOUBLE treatment, except that individuals received an ID number that allows tracing back decisions to individuals after the experiment. These two treatments therefore allow a ceteris-paribus comparison between single- and double-anonymity, which prior field experiments (Cardenas 2014) and the majority of existing lab experiments on this topic have not provided (Barmettler, Fehr, and Zehnder 2012). Second, a further single-anonymity treatment was implemented that involved disclosing the individual decision personally to the experimenter - referred to as REVEAL. This method is not uncommon for lab-in-the-field experiments (e.g. Henrich et al. 2006; Rustagi, Engel, and Kosfeld 2010). Two prior studies have included in their single-anonymity treatment procedures that required subjects to disclose their decisions personally to the experimenter (Cardenas 2014; Hoffman, McCabe, and Smith 1996). Their designs however do not allow to disentangle the effect of personal disclosure from single-anonymity itself. The two single-anonymity treatments in this study allow to do so. Third, besides the common DG, as a measure of pro-social preferences in a non-strategic setting, the Joy of Destruction Mini-Game (JoD) was implemented. This two-person experiment allows subject to engage in anti-social behavior in a dyadic interaction. With its focus on anti-social behavior the JoD presumably induces stronger social EDE than cooperative experiments such as the Prisoner's Dilemma. Prior contributions by Cardenas (2014) and Lesorogol & Ensminger (2014) from the field focus solely on the non-strategic DG.

The results suggest that following a strict double-anonymity procedure is not necessarily needed to minimize social EDE. In both the DG and JoD, behavior in the IDNUM treatment is not significantly different from the double-anonymity procedure (DOUBLE). IDNUM procedures create a sense of anonymity, even though decisions can be linked to subjects through a unique ID after the experimental session. Whether disclosing decisions personally to the experimenter induces additional social EDE cannot be consistently answered by the obtained results. The REVEAL-IDNUM treatment comparison yields significant differences: subjects increase transfers in the DG and are less likely to engage in destructive behavior in the JoD. At the same time, behavior in the REVEAL and DOUBLE treatment is however not significantly different. While these results have to be interpreted with caution, researchers are one the safe side if they design their experiment without the personal disclosure of decisions to the experimenter during the sessions. Yet, full double-anonymity is not required to minimize social EDE.

2. Literature Review

In addition to the two studies that implemented both single and double-anonymous DG in field settings (Cardenas 2014; Lesorogol and Ensminger 2014), a number of lab experiments have investigated the effect of double-anonymity on sharing behavior in the DG[5]. No prior studies have investigated anonymity effects in the strategic context of anti-social behavior. Table 1 summarizes seven experimental studies and their characteristics that include both double-anonymity and single-anonymity treatments in the DG. The dictator game of the present study is also included in the table.

The majority of studies do not find that subject-experimenter anonymity significantly affects behavior: out of seven studies three find that double-anonymity reduces pro-social behavior (Cardenas 2014; Lesorogol and Ensminger 2014; Hoffman et al. 1994). Hoffman et al (1994) indicate that the share of dictators transferring nothing increases from 20% under single-anonymity to more than two thirds under double-anonymity. In Cardenas (2014) double-anonymity reduces average transfers by 22%, i.e. 10% of the total endowment. Lesorogol and Ensminger (2014) report a decrease in average transfers by 30%, which corresponds to 14.5% of the total endowment among their US sample (they do not find significant treatment effects for their Kenyan samples).

All prior studies varied in addition to the experimenter-subject anonymity other aspects of the experimental design between treatments (see Barmettler, Fehr, and Zehnder 2012 for an extended discussion of differences). Most studies vary between treatments the way decisions are taken (Hoffman et al. 1994; Bolton, Katok, and Zwick 1998; Cadsby, Servátka, and Song 2010;

[5] Other types of experiments are less well studied in this regard, potentially because assuring double-anonymity is logistically more demanding if pay-offs depend on the decisions of more than one person. Cox & Deck (2005) find that second movers in the single-anonymous trust game are more likely to act trustworthy as under double-anonymity. Contrary to this, both Deck et al. (2013) and Barmettler et al. (2012) find no evidence for anonymity effects both among first and second movers. Similar, contradicting results exist for the Ultimatum Game: Bolton & Zwick (1995) find evidence for an anonymity effect among second movers; Barmettler et al. (2012) find no behavioral difference under different anonymity conditions. Laury et al. (1995) is the only study focusing on cooperative behavior in a public good experiment. Here, subjects' behavior is not significantly affected by double-anonymity. Two other relevant studies focused on methods beyond lab experiments. Alpizar et al. (2008) conducted a natural experiment by collecting donations for a national park in Costa Rica. They manipulated the degree of anonymity towards the collector, but find no significant difference in donations. In a referendum based contingent valuation study, List et al. (2004) examine to what extent double-anonymity affects stated preferences for donations for an environmental NGO. Using the random response technique to assure double-anonymity significantly reduces the likelihood of voting for a binding donation compared to the conventional non-anonymous elicitation, both in a hypothetical and real setting.

Cardenas 2014). In some cases, details in the experimental instructions (Hoffman et al. 1994), the available action set (Bolton, Katok, and Zwick 1998) or the signing of receipts (Barmettler, Fehr, and Zehnder 2012) are varied as well. Several studies also provide direct information in the instructions that either double anonymity is lacking or assured (Barmettler, Fehr, and Zehnder 2012; Bolton, Katok, and Zwick 1998; Hoffman et al. 1994), which may accentuate anonymity effects. Lastly, two studies involve the direct disclosure of decisions to the experimenter (Hoffman, McCabe, and Smith 1996; Cardenas 2014).

The two studies that find significant anonymity effects are no exception[6]. It is therefore difficult to assert that observed treatment differences are driven indeed by variations of experimenter-subject anonymity and not by other variations in the experimental design or procedures. In the case of Hoffman et al. (1994) and Cardenas (2014) subjects in the double-anonymity treatments are directly endowed with cash, while subjects in the single-anonymity treatments decide either with decision sheets or through a questionnaire. Taking decisions with cash has been shown to increase generosity in the meta-analysis by Engel (2011). As a consequence the anonymity effect may in fact be underestimated. Moreover, Hoffmann et al. (1994) varies the experimental instructions (under double anonymity the endowment is described as "provisional allocation", while in the single anonymity treatment it is not) and explicitly highlights single-anonymity in the respective treatment. Cardenas (2014) implemented the double-anonymity treatments after the single-anonymity treatments in the same village. Behavior in the double-anonymity treatment could be affected by cross-talk and learning effects. Moreover, subjects were directly asked by the experimenter for their transfers in the single-anonymity treatment (Cardenas 2014), which likely induces substantial EDE due to the close interaction between subject and experimenter.

[6] A detailed comparison of single- and double-anonymous procedures in Lesorogol and Ensminger (2014) is not possible, since the experimental protocol is not publicly available.

Table 1. Characteristics of previous experiments on experimenter-subject anonymity in DG.

Authors	Year	Type	N (observations with double-anonymity)	Protocol publicly available	Decision Task		Differences	Significant Effect (p<0.1)
					Single Anonymity	Double Anonymity		
Hoffman et al.[a]	1994	Lab	60 (36)	yes	Decision sheet	Cash	decision task (sheet vs cash), instructions ("provisional allocation" of the endowment under double-anonymity), instructions highlight double-anonymity	yes
Hoffman et al.[b]	1996	Lab	78 (41)	no	Cash with disclosure of decision	Cash	--	no
Bolton et al.[c]	1998	Lab	60 (33)	yes	Cards	Cash	decision task (cards vs cash), different action sets, instructions highlight single-anonymity	no
Cadsby et al[d]	2010	Lab	233 (116)	no	Decision sheet	Cash	decision task (sheet vs cash)	no
Barmettler et al.	2012	Lab	103 (54)	yes	Decision sheet	Decision sheets w. monitor	receipt with pay-off signed only in single-anonymity, instructions highlight double-anonymity	no
Cardenas	2014	Lab-in-the-Field	45 (15)	no	Questionnaire	Cash	decision task (questionnaire vs cash), DB treatments run after single-anonymity on one of two villages covered by single-anonymity treatment	yes
Lesorogol & Ensminger	2014	Lab-in-the-Field	217 (68) split between 2 Kenyan and 1 US sample	only double-anonymity	*na*	Cash	*na*	only for US sample
This study	--	Lab-in-the-Field	240 (80)	--	a) Cash with disclosure of decision, b) Cash	Cash	One single-anonymity treatment with ID numbers to link decisions to individuals, one single-anonymity treatment with personal disclosure of decisions	mixed

[a] based on the "FHSS Results Divide 10$ Dictator" and "Double Blind 1" treatment, since behaviour is not compared between the FHSS and "Double Blind 2" treatment. [b] based on comparison between DB2 and SB1, since these treatment are the most similar. [c] based on the "1Game6" and "ANON" treatment. [d] "Single Blind – single sex" and double blind – single sex" treatment.

The present study tries to avoid these pitfalls of prior studies by not altering aspects of the experimental design that may confound the analysis of anonymity effects. In the DG, decisions in all treatments are taken with cash and differences in the instructions were minimized as much as possible. One single-anonymous treatment is designed to resemble exactly the double-anonymous procedures except the use of ID numbers that allow to identify individual decisions. Another single-anonymity treatment is designed to include the personal disclosure of decisions to the experimenter. Two prior studies included similar procedures in their experiments, but they do not allow to disentangle the effect of personal disclosure from single-anonymity as such (Hoffman, McCabe, and Smith 1996; Cardenas 2014). The present study allows this by comparing two single-anonymity treatments that either include or exclude the personal disclosure of decisions. The experimental design and procedures are discussed in the following section in more detail.

3. Experimental Design and Procedures

3.1. Choice of Experiments

As mentioned above, this paper applies different treatments that vary the degree of subject-experimenter anonymity in the DG and JoD. The DG was chosen, as it is likely to induce relatively strong social EDE due to the non-strategic decision subjects take. Social norms have been shown to be a powerful driver of behavior in the DG (Krupka and Weber 2013) and subjects are likely aware of the social-desirable behavior. Double-anonymity is consequently likely to reduce social EDE in this experiment and the estimated effects could be considered as an upper bound for other experiments. Additionally, the majority of former lab and field studies implement double-anonymity in the DG (see Section 2), thus allowing for an easy comparison of findings.

The choice of the JoD was motivated to study double-anonymity and social EDE in a strategic setting. Behavior in many cooperative experiments such as the Prisoner's Dilemma or Public Good Games, however, has been shown to be highly conditional on the decision of others (e.g. Fischbacher, Gächter, and Fehr 2001). Hence, social norms are likely less salient and less influential in these types of experiments. The JoD, in contrast, allows for anti-social or destructive behavior, which is likely perceived as socially inappropriate independently of the decision of the other player[7]. Hence, I expect the JoD to induce stronger social EDE than cooperative experiments do and potentially larger anonymity effects. In particular, social EDE are assumed to prescribe non-destructive behavior.

[7] Similar to norms of conditional cooperation, there might be a conditional social norm that destruction is socially appropriate, if the other player chooses "destroy" as well. To my knowledge, this has not been studied empirically. Rabin (1993) theoretically shows that fairness motives can results in mutually detrimental equilibria (corresponding to both players choosing "destroy" in the JoD).

3.2. Dictator Game

In the well-known Dictator Game (DG), pioneered by Forsythe et al. (1994), one subject receives a fixed endowment and has to decide how much to allocate to her partner. The second subject benefits from the transfer, but can neither accept nor reject it. In each session ten subjects acted as senders, who each received an envelope with 60 Namibian Dollar (ND) in cash (≈ 9 USD PPP). The receivers participated in two later sessions. It was known to the senders that their identity will not be revealed to their matched partner at any point. The endowment was given to senders in four bills (2 × 10 ND, 2 × 20 ND) allowing division of the transfer money to the second player in 10 ND increments[8].

Table 2. Pay-Off Structure JoD.

		Player B	
		Destroy	*Not Destroy*
Player A	*Destroy*	8, 8	18, 10
	Not Destroy	10, 18	20, 20

3.3. Joy of Destruction Mini-Game

In the Joy-of-Destruction (JoD) Mini-Game, developed by Abbink & Herrmann (2011), individuals decide whether to engage in destructive behavior. Subjects were randomly matched with someone in the same session (n=10), without knowing the partner's identity. Each subject received 20 points (≈ 3 USD PPP) and made one decision: whether to destroy 10 points of the partner's endowment at a cost of 2 points. The four possible outcomes are summarized in Table 2. Each subject makes the decision without knowing their partner's decision. The strictly dominant strategy of a rational, pay-off maximizing individual would be “not destroy”. Since destruction does not provide any material benefits for the destroying party and the game is anonymous and not repeated, selfish, fairness or reciprocity are ruled out as possible motives. Destructive behavior is hence most likely attributed to spite or nastiness (Abbink and Sadrieh 2009).

3.4. Treatments and Experimental Procedures

This section contains a brief description of the three different treatments. A detailed description of the procedures and protocols, as well as a general overview of different techniques to assure double-anonymity, can be found in the Supplementary Online Material (SOM), Appendix A-E.

DOUBLE: In the DG, unmarked envelopes - that were picked by subjects themselves - assured double-anonymity. The symmetric design of the JoD required subjects to use keys and boxes to deposit their decision sheets and collect pay-offs. Subjects individually drew envelopes that contained a unique key with

[8] Heavy manila envelopes were used to prevent subjects and experimenters from seeing whether or how much money was put in the envelopes.

access to one box. A set of second keys were used by the experimenter to access the decision sheets and deposit the pay-offs in the boxes.

IDNUM: This treatment closely resembled the DOUBLE treatment, except the use of ID numbers that allow to relate decisions to individuals after the session.

REVEAL: Here, subjects disclosed their decision personally, but in private, to one of the experimenters. In the DG, subjects showed their unsealed envelope with the money they decided to transfer to one experimenter in private. In the JoD, boxes and keys were not used. Subjects showed the decision sheet to one of the experimenters in private after taking the decision. Pay-offs were distributed to subjects in private at the end of each session.

While the procedures in the DG and JoD differ, both experiments share several characteristics. In all three treatments, the actual decision-making was identical (in private inside the booth, with cash in the DG and decision sheets in the JoD). The respective procedures were explained to the subjects in detail before the decision-making, so the degree of anonymity was known by subjects. However, in none of the experiments was the actual degree of experimenter-subject anonymity explicitly highlighted. After the experiment an individual questionnaire was administered for all three treatments. Note that individual decisions can only be related to survey information in the IDNUM and REVEAL treatment. In addition to the pay-offs, subjects received a show-up fee of 10 ND (≈ 1.5 USD PPP) in the DG and 20 ND (≈ 3 USD PPP) in the JoD. Average earnings are 63 ND (≈ 9.5 USD PPP) for the DG and 44 ND (≈ 6.6 USD PPP) in the JoD.

In each sampled village (see Section 3.5) two DG and two JoD sessions were implemented on the same day. After two experimental sessions in the morning that included only dictators of the DG, the corresponding receivers of the DG participated in two afternoon sessions, who also played the JoD[9]. Here, subjects individually drew one of the envelopes with the transfers from a bag after the JoD was finalized[10]. All experiments in one village were conducted on the same day. To minimize cross talk, the sessions of the same experiment (but with two different treatments) were conducted consecutively. Both experiments were implemented as a between-subject design, i.e. each subject received only one treatment, to minimize potential between-games EDE. The team of research assistants (experimenters) and their respective roles in the sessions was not changed throughout data collection to hold experimenter effects constant. Each village was randomly assigned a combination of two different treatments for both the DG and JoD[11].

[9] The senders did not receive the information that receivers will participate in another experiment.

[10] The receivers were not aware that they will receive transfers from the DG until the JoD was finalized to assure that potential expectations of the transfer will not influence decisions in the JoD.

[11] Three treatments yield 12 unique combinations of two different treatments, considering the order of the treatments. These twelve combinations were randomly assigned to the villages for the DG and JoD independently. The treatment plan can be found in the SOM, Appendix D.

3.5. Sampling

The experiments were conducted in 12 different villages in the Kavango East Region of Namibia. The design was pre-tested in two additional villages during seven sessions (4 DG, 3 JoD). The experimental protocols were translated from English into the local language and back-translated by two different research assistants. Conflicting and ambiguous parts were then jointly changed. In order to cover a variety of contexts, villages were selected along two roads: to the east of the regional capital along the Kavango river in a relatively densely populated area and to the south of the regional capital in a more sparsely populated area. In each village two sessions each of the DG and JoD were conducted.

Subjects were randomly selected at the village level during a village meeting that was announced by the respective village headmen few days earlier. On the day of the experiment and after a general introduction of the research team, each present adult (above 18 years) participated in a lottery that determined whether and in which session they would participate. The final DG and JoD datasets contain 240 and 237 observations respectively[12]. Three observations of the JoD had to be excluded due to missing decisions. The socio-economic characteristics of the DG and JoD sample can be found in the SOM. Joint F-tests confirm that randomization ensured the absence of significant differences between treatment groups in terms of observable socio-economic characteristics (see SOM, Appendix F)[13].

4. Results

4.1. Dictator Game

Table 3 summarizes the mean transfer of dictators to the receiver in the DG and the share of subjects who did not transfer anything by treatment. Overall, subjects sent on average 6.2 ND (10.3 %) of their endowment to their partners. Transfers are on average highest in the REVEAL treatment. Here, also more than half of the participants sent 10 ND or more to their partners. Figure 1 illustrates the distribution of decisions by treatment. In the REVEAL treatment, participants more frequently transfer nothing and less frequently 10 or 20 ND.

[12] The dataset and code of the analysis is available from the author upon request.

[13] There are significant differences for age and education in the JoD Sample. The presented results are robust to the inclusion of these controls in the analysis (see SOM, Appendix G). Observations from the DOUBLE treatment are not included in this analysis, since decisions cannot be related to personal information from the post-experiment questionnaire in this treatment.

Table 3. Observations and Decisions by Treatment – DG.

Treatment	Observations	Mean Transfer	% > 0
DOUBLE	80	5.38	40.0
IDNUM	80	5.25	35.0
REVEAL	80	7.88	52.5

Table 4 includes three different models with the DG decisions as dependent variable. Due to the assignment of treatments at the village level, standard errors are clustered in all following regression analyses at the village level (Abadie et al. 2017)[14]. OLS and Tobit regressions indicate that the IDNUM treatment results in significantly lower dictator transfers compared to the REVEAL treatment. Also, subjects in the IDNUM treatment are 17.5% less likely to transfer positive amounts to the second player as indicated by Model 3. The DOUBLE treatment has no significant effect on transfers relative to the REVEAL treatment, even though the effect size is similar to the IDNUM treatment. At the same time, the IDNUM and DOUBLE treatment effects are not significantly different in all three models as indicated by Wald tests.[15]

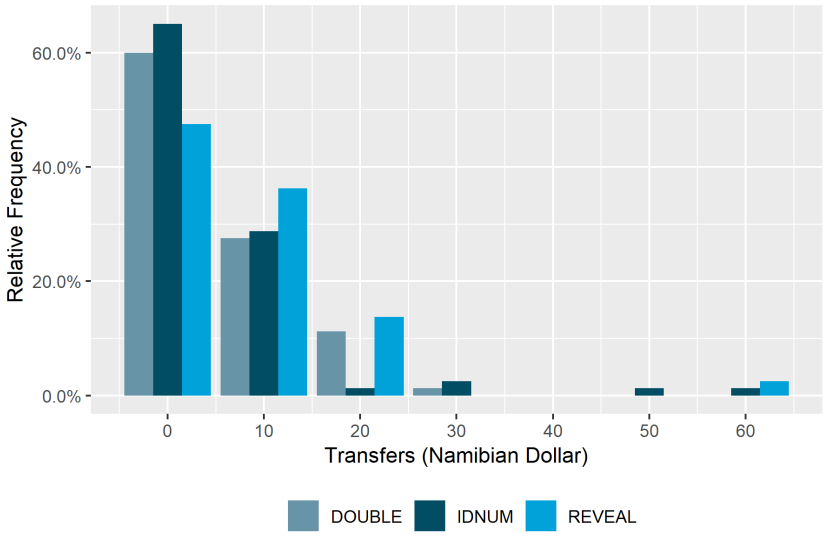

Figure 1. Dictator Game Transfers by Treatment.

[14] Due to the low number of clusters (the effective number of clusters is 8, see Carter, Schnepel, and Steigerwald 2017), p-values derived with the wild cluster bootstrap method are reported in addition for the OLS and Probit models. This method performs relatively well with few clusters (Cameron, Gelbach, and Miller 2008).

[15] In the post-experiment questionnaire enumerators furthermore directly asked for the individual decisions (see SOM, Appendix H). If the DOUBLE treatment reduced social EDE successfully, one would expect differences between stated and real decisions in this treatment. Subjects, whose giving was reduced by the DOUBLE (relative to the REVEAL) treatment, would be expected to inflate their self-reported decisions to conceal their actual (more selfish) behavior. The difference between real and stated decisions for the DOUBLE treatment is statistically significant ($p < 0.05$), which is not the case for the IDNUM treatment ($p < 0.3$). These results therefore contrast the regression results.

Finding 1: Pro-social behavior is less likely under the IDNUM compared to REVEAL treatment. At the same time, behavior in the DOUBLE treatment is not significantly different from the behavior observed in the REVEAL treatment.

Finding 2: There is no significant behavioral difference between DOUBLE and IDNUM procedures in the DG.

Table 4. Regression Results Dictator Game.

	(1)		(2)	(3)	
	OLS	WCB	Tobit	Probit	WCB
DV	Transfer (ND)			Transfer > 0 ND	
DOUBLE	-2.500	0.249	-5.907	-0.122	0.312
	(-1.59)		(-1.35)	(-1.18)	
IDNUM	-2.625*	0.098	-7.210**	-0.172*	0.085
	(-1.82)		(-2.05)	(-1.93)	
N	240		240	240	
p > Wald chi²: DOUBLE=IDNUM	0.94	0.942	0.76	0.60	0.656

Model 3 reports marginal effects; t statistics in parentheses, Bootstrapped clustered standard errors at the village level (R=500), WCB columns show the p-values derived from the wild cluster bootstrap estimation (Cameron, Gelbach, and Miller 2008)
** p < 0.10, ** p < 0.05, *** p < 0.01*

4.2. Joy of Destruction Game

The experimental procedures for the DOUBLE and IDNUM treatment involved opening and closing locked boxes (see SOM, Appendix E.3). Older and less-educated participants are significantly less likely to manage this task independently (see SOM, Appendix I). Out of 157 subjects, 40 required assistance (circa 25.5 %), which was provided by one experimenter inside the booth. In the DOUBLE treatment, assistance lifted strict double-anonymity due to the experimenter observing the box, where the decision sheets was placed in (even if the experimenter might not observe the decision personally). In many cases, however, subjects failed to place the decision sheets inside the envelope, allowing the experimenter to directly see the decision. In case subjects – who have been either assigned to the DOUBLE or IDNUM treatment – received assistance, observations are therefore coded as treated with REVEAL. In the following analysis, there is consequently a distinction between treatment as originally assigned and eventually given to subjects.

Figure 2 shows the relative frequencies of destructive behavior ("destroy" decision) by assigned treatments and treatments given (treated). Destructive behavior is most frequently observed in the IDNUM treatment (assigned and treated both 29.1%), followed by the DOUBLE treatment (assigned: 18%, treated:

19.4%). Under single-anonymity the smallest fraction of subjects choose "destroy" (assigned: 12.5%, treated: 15.8%).[16]

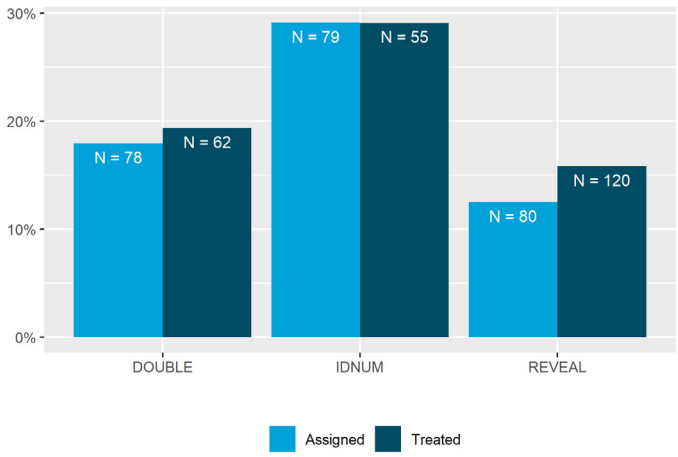

Figure 2. Relative frequency of destructive behavior by treatments assigned and treated.

Due to the one-sided non-compliance of subjects who were assigned to the IDNUM or DOUBLE treatment but were treated with REVEAL, the intention-to-treat (ITT) effect and the average treatment effect on the treated (ATT) is estimated separately (see Table 5). The ATT is estimated by instrumenting the treatment delivered by the exogenous (random) treatment assignment (Angrist 2006). This allows to control for the non-random self-selection of subjects into the delivered treatments; as noted above older and less-educated subjects were more likely to require assistance in the DOUBLE and IDNUM treatment and thus more likely to be actually treated with SINGLE. Similar to the results in the DG, destructive behavior is significantly more likely under the IDNUM compared to the REVEAL treatment. Subjects assigned to the IDNUM treatment are 16% more likely to engage in destructive behavior, while subject treated with IDNUM are 23% more likely to act destructively. The ITT and ATT effect for the DOUBLE treatment are smaller in magnitude and not significantly different from zero. The estimated IDNUM and DOUBLE treatments effects are however for both specifications not significantly different as indicated by Wald tests.[17]

---

[16] At the village level the frequency of destructive behavior in the JoD and the average transfer in the DG is positively correlated (Pearson's correlation coefficient: 0.51, $p < 0.09$). Due to the low number of observations for village level outcomes, conclusions have to be drawn with caution. However, these results echo findings of Basurto et al. (2016), who report that high levels of pro- and anti-social behavior coexist in villages inside Marine Protected Areas in Mexico. At the individual level, Sadrieh and Schröder (2016) report a positive correlation between pro- and anti-social behavior among on third of their sample. Very few subjects behave anti-socially without exhibiting pro-social behavior. Similar findings are reported by Zizzo and Fleming (2011), who find a positive correlation between giving in a dictator game and anti-social behavior in a money burning game.

[17] Individual decisions of the JoD were also elicited in the post-experiment questionnaire (see SOM, Appendix H). In contrast to the DG, both real and stated decisions are similar within the treatments and these differences are found to be not statistically significant for each of the three treatments.

Table 5. Regression Results – JoD.

	(1) Probit (ITT)	WCB	(2) IV Probit (ATT)
DV	Destructive Behavior		
DOUBLE	0.063	0.459	0.0751
	(0.59)		(0.70)
IDNUM	0.162*	0.083	0.227**
	(1.80)		(2.26)
N	237		237
p > Wald chi²: DOUBLE=IDNUM	0.24	0.186	0.22

Model 1 and 2 report marginal effects; t statistics in parentheses, Bootstrapped clustered standard errors at the village level (R=500), WCB column shows the p-values derived from the wild cluster bootstrap estimation (Cameron, Gelbach, and Miller 2008)
** p < 0.10, ** p < 0.05, *** p < 0.01*

Finding 3: Anti-social behavior is significantly more likely under the IDNUM than REVEAL treatment. At the same time, the DOUBLE treatment does not significantly increase destructive behavior compared to the REVEAL treatment.

Finding 4: There is no significant behavioral difference between DOUBLE and IDNUM procedures in the JoD.

5. Discussion

In this paper, three treatments were introduced that varied the interaction and distance between experimenters and subjects during the decision-making to test whether they induce different levels of social EDE[18]. The starkest treatment, REVEAL, involved disclosed decisions personally to the experimenter. In both the non-strategic DG and the strategic JoD, subjects' behavior in the REVEAL treatment is significantly different from the IDNUM treatment that does not guarantee complete double-anonymity. Decreasing the distance between experimenters and subjects during the decision-making and hence increasing social pressure induces additional EDE, so that subjects act more pro-socially (Finding 1) and less anti-socially respectively (Finding 3). While the comparisons between the REVEAL and IDNUM treatment are statistically significant, the comparisons between the REVEAL and DOUBLE treatment is not statistically significant (even though the

[18] The average transfer rates in the DG are below the average of similar lab-in-the-field experiments (Cardenas and Carpenter 2008) and dictator games in general (Engel 2011). Comparing the distribution of transfers with the distribution synthesized from lab experiments by Engel (2011) highlights that subjects from this study more frequently transfer nothing (57.5% vs 36%). Differences are even more pronounced for subjects who transfer half of their endowment (lab: 16.7%, this study: 1.3%). In comparison with a previous JoD Experiment in Namibia, rates of anti-social behavior are also relatively low (Prediger, Vollan, and Herrmann 2014).

average behavior points towards the same direction). These findings are puzzling, since the DOUBLE treatment provides an even higher degree of anonymity than the IDNUM treatment does. One would consequently expect to observe treatment differences that are at least as strong. The DOUBLE treatment with complete experimenter-subject anonymity may have actually achieved the opposite of its initial objective. For example, the lack of ID numbers in the DOUBLE treatment may have decreased the credibility that subjects' decisions are pay-off relevant or raised general skepticism regarding the true objective of the research [19]. Considering the low number of villages, the level at which treatments were randomly allocated, statistical power remains a concern and potentially explains these contradicting results. Hence, whether the personal disclosure of individual decisions to the experimenter increases social EDE cannot be answered definitely. Considering that methodological questions should be approached conservatively, it is recommended to avoid that subjects personally reveal their decisions to the experimenter during the sessions. At the same time, experimental findings from the field that involve such a personal disclosure of decisions to experimenters have to be interpreted with caution.

The comparisons between the DOUBLE and IDNUM treatment suggest that double-anonymity alone does not reveal significant different behavior (Finding 2 & 4). This is in line with two previous lab studies that also provide nearly perfect ceteris-paribus comparisons. These studies find no significant behavioral difference between single- and double-anonymity procedures (Barmettler, Fehr, and Zehnder 2012; Deck, Servátka, and Tucker 2013)[20]. A fundamental, positive implication for lab-in-the-field experiments is that ID numbers, linking individual decisions to socio-economic survey data, alone, do not induce *additional* social EDE. Consequently, full double-anonymous procedures are not necessarily required to minimize social EDE. The use of ID number provides sufficient privacy for subjects during the experimental sessions, but allow to relate experimental decisions to personal information collected during the post-experiment questionnaire.

The inclusion of two types of experiment - the non-strategic DG and the strategic JoD - was motivated to include an experiment where pay-offs are interdependent. Behavior in both experiments is not easily comparable as the DG elicits pro- and the JoD anti-social behavior and more sophisticated methods are required to assure double-anonymity in the JoD (i.e. locked boxes). Nonetheless, the results from both experiments agree qualitatively, providing further robustness for the findings.

How do these results relate to previous evidence from the field? As noted above, Cardenas (2014) cannot disentangle whether the treatment effect is driven by the way decisions were taken or double-anonymity as such. The presented results suggest that disclosing decisions personally to the experimenter induce additional social EDE, while double-anonymity alone does not affect social EDE relative to single-anonymity.

---

[19] The absence of ID number may have also irritated subjects and resulted in erratic behavior. Yet, the variance of transfers is substantially lower in the DOUBLE treatment (DOUBLE 55.55 ND, IDNUM 106.27 ND, REVEAL 120.74 ND). None of the pairwise comparisons is statistically significant (Mann-Whitney-U Tests).

[20] The design of this study does not allow to infer the absolute level of social EDE but allows to assess relative levels of social EDE under different degrees of anonymity. De Quidt et al. (2018) systematically induced social EDE in an online experiment through manipulating the instructions and find support for EDE of modest size. To what extent these findings can be replicated in lab-in-the-field experiments remains to be studied.

Clearly, the sample has specific characteristics that reduce the generalizability of the findings. Better educated subjects may be more attentive to small procedural differences (such as the use of IDs) and thus may require strict double-anonymous procedures to effectively minimize social EDE. Lesorogol and Ensminger (2014) compare how samples in Kenya and the USA react to double-anonymity. They find that double-anonymity significantly reduces pro-social behavior in the USA, but not in Kenya. This difference may be explained by educational differences of the samples, even though other explanations cannot be ruled out (e.g. cultural). Further evidence from the field with better educated subjects is required to assess how education interacts with experimental procedures that vary the distance during the decision-making and experimenter-subject anonymity.

One major concern is, the substantial share of subjects that did not manage to follow the double-anonymity procedures in the JoD (almost 25%). This procedure seems too complex, especially for older and less-educated subjects. In the JoD, a relatively complex procedure with boxes and locks was also applied in the IDNUM treatment to allow for a ceteris-paribus comparison with the DOUBLE treatment. Potentially, sufficient distance between subjects and experimenters could be also realized through a simpler procedure (using IDs and assigning two different experimenters for the pay-off preparation and distribution). Whether such a procedure is effective in reducing social EDE remains to be tested experimentally.

Lastly, the presented results strongly encourage a clear and consistent documentation of experimental sessions in the field. Experimenters should keep records if subjects require assistance (e.g. older ones) and control for it in the data analysis. More generally, lab-in-the-field experiments would increase their replicability, if information concerning subject-subject and experimenter-subject anonymity, as well as information on the decision-making environment (e.g. in private or public), is provided upon publication.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge. 2017. "When Should You Adjust Standard Errors for Clustering?" Working Paper 24003. National Bureau of Economic Research. https://doi.org/10.3386/w24003.

Abbink, Klaus, and Benedikt Herrmann. 2011. "The Moral Costs of Nastiness." *Economic Inquiry* 49 (2): 631–33. https://doi.org/10.1111/j.1465-7295.2010.00309.x.

Abbink, Klaus, and Abdolkarim Sadrieh. 2009. "The Pleasure of Being Nasty." *Economics Letters* 105 (3): 306–8. https://doi.org/10.1016/j.econlet.2009.08.024.

Alpizar, Francisco, Fredrik Carlsson, and Olof Johansson-Stenman. 2008. "Anonymity, Reciprocity, and Conformity: Evidence from Voluntary Contributions to a National Park in Costa Rica." *Journal of Public Economics* 92 (5–6): 1047–60. https://doi.org/10.1016/j.jpubeco.2007.11.004.

Angrist, Joshua D. 2006. "Instrumental Variables Methods in Experimental Criminological Research: What, Why and How." *Journal of Experimental Criminology* 2 (1): 23–44. https://doi.org/10.1007/s11292-005-5126-x.

Barmettler, Franziska, Ernst Fehr, and Christian Zehnder. 2012. "Big Experimenter Is Watching You! Anonymity and Prosocial Behavior in the Laboratory." *Games and Economic Behavior* 75 (1): 17–34. https://doi.org/10.1016/j.geb.2011.09.003.

Basurto, Xavier, Esther Blanco, Mateja Nenadovic, and Björn Vollan. 2016. "Integrating Simultaneous Prosocial and Antisocial Behavior into Theories of Collective Action." *Science Advances* 2 (3): e1501220. https://doi.org/10.1126/sciadv.1501220.

Bolton, Gary E., Elena Katok, and Rami Zwick. 1998. “Dictator Game Giving: Rules of Fairness versus Acts of Kindness.” *International Journal of Game Theory* 27 (2): 269–99. https://doi.org/10.1007/s001820050072.

Bolton, Gary E., and Rami Zwick. 1995. “Anonymity versus Punishment in Ultimatum Bargaining.” *Games and Economic Behavior* 10 (1): 95–121. https://doi.org/10.1006/game.1995.1026.

Cadsby, C. Bram, Maroš Servátka, and Fei Song. 2010. “Gender and Generosity: Does Degree of Anonymity or Group Gender Composition Matter?” *Experimental Economics* 13 (3): 299–308. https://doi.org/10.1007/s10683-010-9242-8.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. “Bootstrap-Based Improvements for Inference with Clustered Errors.” *The Review of Economics and Statistics* 90 (3): 414–27.

Cardenas, Juan Camilo. 2014. “Social Preferences Among the People of Sanquianga in Colombia.” In *Experimenting with Social Norms: Fairness and Punishment in Cross-Cultural Perspective*, edited by Jean Ensminger and Joseph Henrich. Russell Sage Foundation.

Cardenas, Juan Camilo, and Jeffrey Carpenter. 2008. “Behavioural Development Economics: Lessons from Field Labs in the Developing World.” *The Journal of Development Studies* 44 (3): 311–38. https://doi.org/10.1080/00220380701848327.

Carter, Andrew V., Kevin T. Schnepel, and Douglas G. Steigerwald. 2017. “Asymptotic Behavior of a T-Test Robust to Cluster Heterogeneity.” *The Review of Economics and Statistics* 99 (4): 698–709. https://doi.org/10.1162/REST_a_00639.

Cassar, Alessandra, Andrew Healy, and Carl von Kessler. 2017. “Trust, Risk, and Time Preferences After a Natural Disaster: Experimental Evidence from Thailand.” *World Development* 94 (June): 90–105. https://doi.org/10.1016/j.worlddev.2016.12.042.

Chakravarty, Surajeet, Miguel A. Fonseca, Sudeep Ghosh, and Sugata Marjit. 2016. “Religious Fragmentation, Social Identity and Cooperation: Evidence from an Artefactual Field Experiment in India.” *European Economic Review*, Social identity and discrimination, 90 (November): 265–79. https://doi.org/10.1016/j.euroecorev.2015.12.006.

Cilliers, Jacobus, Oeindrila Dube, and Bilal Siddiqi. 2015. “The White-Man Effect: How Foreigner Presence Affects Behavior in Experiments.” *Journal of Economic Behavior & Organization*, Economic Experiments in Developing Countries, 118 (October): 397–414. https://doi.org/10.1016/j.jebo.2015.03.015.

Cox, James C., and Cary A. Deck. 2005. “On the Nature of Reciprocal Motives.” *Economic Inquiry* 43 (3): 623–35. https://doi.org/10.1093/ei/cbi043.

Crowne, Douglas P., and David Marlowe. 1960. “A New Scale of Social Desirability Independent of Psychopathology.” *Journal of Consulting Psychology* 24 (4): 349–54. https://doi.org/10.1037/h0047358.

Deck, Cary, Maroš Servátka, and Steven Tucker. 2013. “An Examination of the Effect of Messages on Cooperation under Double-Blind and Single-Blind Payoff Procedures.” *Experimental Economics* 16 (4): 597–607. https://doi.org/10.1007/s10683-013-9353-0.

Engel, Christoph. 2011. “Dictator Games: A Meta Study.” *Experimental Economics* 14 (4): 583–610. https://doi.org/10.1007/s10683-011-9283-7.

Fischbacher, Urs, Simon Gächter, and Ernst Fehr. 2001. “Are People Conditionally Cooperative? Evidence from a Public Goods Experiment.” *Economics Letters* 71 (3): 397–404. https://doi.org/10.1016/S0165-1765(01)00394-9.

Forsythe, Robert, Joel L. Horowitz, N. E. Savin, and Martin Sefton. 1994. “Fairness in Simple Bargaining Experiments.” *Games and Economic Behavior* 6 (3): 347–69. https://doi.org/10.1006/game.1994.1021.

Henrich, Joseph, Richard McElreath, Abigail Barr, Jean Ensminger, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, et al. 2006. “Costly Punishment Across Human Societies.” *Science* 312 (5781): 1767–70. https://doi.org/10.1126/science.1127333.

Hoffman, Elizabeth, Kevin McCabe, Shachat Keith, and Vernon Smith. 1994. “Preferences, Property Rights, and Anonymity in Bargaining Games.” *Games and Economic Behavior* 7 (3): 346–80.

Hoffman, Elizabeth, Kevin McCabe, and Vernon L. Smith. 1996. “Social Distance and Other-Regarding Behavior in Dictator Games.” *The American Economic Review* 86 (3): 653–60. https://doi.org/10.2307/2118218.

Krupka, Erin L., and Roberto A. Weber. 2013. “Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?” *Journal of the European Economic Association* 11 (3): 495–524. https://doi.org/10.1111/jeea.12006.

Laury, Susan K., James M. Walker, and Arlington W. Williams. 1995. "Anonymity and the Voluntary Provision of Public Goods." *Journal of Economic Behavior & Organization* 27 (3): 365–80. https://doi.org/10.1016/0167-2681(94)00073-N.

Lesorogol, Carolyn K., and Jean Ensminger. 2014. "Double-Blind Dictator Games in Africa and the United States: Differential Experimenter Effects." In *Experimenting with Social Norms: Fairness and Punishment in Cross-Cultural Perspective*, edited by Jean Ensminger and Joseph Henrich. Russell Sage Foundation.

Levitt, Steven D, and John A List. 2011. "Was There Really a Hawthorne Effect at the Hawthorne Plant? An Analysis of the Original Illumination Experiments." *American Economic Journal: Applied Economics* 3 (1): 224–38. https://doi.org/10.1257/app.3.1.224.

List, John A., Robert P. Berrens, Alok K. Bohara, and Joe Kerkvliet. 2004. "Examining the Role of Social Isolation on Stated Preferences." *The American Economic Review* 94 (3): 741–52. https://doi.org/10.2307/3592951.

Millham, Jim, and Richard W Kellogg. 1980. "Need for Social Approval: Impression Management or Self-Deception?" *Journal of Research in Personality* 14 (4): 445–57. https://doi.org/10.1016/0092-6566(80)90003-3.

Nederhof, Anton J. 1985. "Methods of Coping with Social Desirability Bias: A Review." *European Journal of Social Psychology* 15 (3): 263–80. https://doi.org/10.1002/ejsp.2420150303.

Prediger, Sebastian, Björn Vollan, and Benedikt Herrmann. 2014. "Resource Scarcity and Antisocial Behavior." *Journal of Public Economics* 119 (November): 1–9. https://doi.org/10.1016/j.jpubeco.2014.07.007.

Quidt, Jonathan de, Johannes Haushofer, and Christopher Roth. 2018. "Measuring and Bounding Experimenter Demand." *American Economic Review* 108 (11): 3266–3302. https://doi.org/10.1257/aer.20171330.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *The American Economic Review* 83 (5): 1281–1302.

Rustagi, Devesh, Stefanie Engel, and Michael Kosfeld. 2010. "Conditional Cooperation and Costly Monitoring Explain Success in Forest Commons Management." *Science* 330 (6006): 961–65. https://doi.org/10.1126/science.1193649.

Sadrieh, Abdolkarim, and Marina Schröder. 2016. "Materialistic, pro-Social, Anti-Social, or Mixed – A within-Subject Examination of Self- and Other-Regarding Preferences." *Journal of Behavioral and Experimental Economics* 63 (August): 114–24. https://doi.org/10.1016/j.socec.2016.05.009.

Warner, Stanley L. 1965. "Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias." *Journal of the American Statistical Association* 60 (309): 63–69. https://doi.org/10.1080/01621459.1965.10480775.

Zizzo, Daniel John. 2010. "Experimenter Demand Effects in Economic Experiments." *Experimental Economics* 13 (1): 75–98. https://doi.org/10.1007/s10683-009-9230-z.

Zizzo, Daniel John, and Piers Fleming. 2011. "Can Experimental Measures of Sensitivity to Social Pressure Predict Public Good Contribution?" *Economics Letters* 111 (3): 239–42. https://doi.org/10.1016/j.econlet.2011.02.021.