UC Santa Barbara
Departmental Working Papers

Title
Social and Moral Norms in the Laboratory

Permalink
https://escholarship.org/uc/item/6rv7x0tf

Authors
Charness, Gary
Schram, Arthur

Publication Date
2012-02-05

Social and Moral Norms in the Laboratory

Arthur Schram* and Gary Charness[+]

February 5, 2012

Abstract: Social norms involve observation by others and external sanctions for violations, while moral norms involve introspection and internal sanctions. We develop a simple model of individual preferences that incorporates moral and social norms. We then examine dictator choices, where we create a shared understanding by providing advice from peers with no financial payoff at stake. We vary whether advice is given, as well as whether choices are made public. This design allows us to explicitly separate the effects of moral and social norms. We find that choices are in fact affected by a combination of observability and the shared understanding.

* CREED, Amsterdam School of Economics, Roetersstraat 11, 1018 WB Amsterdam. Email: Schram@uva.nl.

[+] Department of Economics, University of California at Santa Barbara, 2127 North Hall, Santa Barbara, California, 93106-9210. Email: charness@econ.ucsb.edu.

1. Introduction

Economic theory typically presumes that people know their own preferences and make decisions on the basis of these preferences. Yet in many situations people rely on social and moral norms to guide their behavior, so that one's utility appears to depend on the preferences or the degree of one's adherence to the perceived norms of others. In this study, we present a simple model of individual preferences that allows for this dependence. We then develop a tool to create norms in the lab and to distinguish between distinct types of norms. We consider an allocation decision (featuring six possible alternative choices) to be made amongst three people, after advice from other people who have no financial payoff at stake. We also vary whether the chosen action (but not the advice given) is observable to the other people in the experiment.

To distinguish 'social' norms from other types, Elster (2009) offers a categorization of types of norms that makes explicit the importance of sociality. He defines a social norm to be a "non-outcome-oriented injunction to act or to abstain from acting, sustained by the sanctions that others apply to norm violators." Sanctions need not be monetary, but could instead involve shame. Thus, *observation* is a critical element of social norms. Social norms are powerful due to the common desire for approval from one's peers or colleagues, and observation brings this issue to the fore. One's identity (see Akerlof and Kranton, 2000) is shaped in part by this desire, and one acts in accord with the social norm in order to be a respected group member.[1]

On the other hand, in Elster's view moral norms have been internalized, so that 'punishment' comes from within and is manifested by a negative view of one's self, perhaps involving emotions such as guilt or remorse.[2] Moral norms require introspection rather than external observation. Both require a shared understanding about what one ought to do. While

---

[1] See Charness, Rigotti, and Rustichini (2007) and Andreoni and Bernheim (2009) for some experimental evidence concerning the effect of observation.
[2] Note that 'internalized social norms' may be seen as a special case of moral norms.

moral norms involve a decision maker internalizing this shared understanding through an emotional mechanism like guilt, social norms involve an emotional mechanism (like shame) that induces a decision maker to adhere to the shared understanding when others observe her decision. Traditionally, much focus has been on social norms. Note however, that an increasing number of decisions are being made behind a computer screen. This means that decisions are increasingly often being made privately, hence moral norms are becoming more important.

Elster's definition focuses directly on the role that norms play in the individual decision-making process. Others have attempted to define norms based on the strategic interaction between individuals and to derive norms indirectly from observed behavior (see Reuben and Riedl 2011 for a discussion).[3] The advantage of Elster's definition over these is that it focuses directly on individual decision rules, without any need for assumptions about the strategic interaction between individuals.

In this regard, a major contribution of our study is that we use a simple design that allows us to focus directly on such decision rules. Not only does it allow us to measure the norm that is applicable in a specific situation, it is also informative of the specific behavior that the norm prescribes. One may think of decision rules as expressions of the prescriptions that arise from the norm. Studying the decision rules directly measures the imperative associated with the norm. As an example, while Krupka and Weber (2010) use a technique to try to identify which action is most appropriate, they are not able to explain how this is translated into a decision rule.[4]

---

[3] For example, Young (2008) sees norms (which he calls 'conventions') as the equilibria of an underlying game and applies the term only to games with multiple equilibria; norms cannot enforce non-equilibrium behavior (other authors that see norms as equilibria to games include Binmore and Samuelson 1994, Gintis 2009, and Ellingson 2011). On the other hand, Bicchieri (2006) sees conventions as a coordination device, followed out of individual self-interest. Like Elster (2009), Bicchieri contrasts conventions with social norms, which in her view enforce out-of-equilibrium behavior in situations where individual optimization is inefficient.
[4] We would like to thank Erin Krupka for her comments in this regard.

Of course, we are not the first in economics to consider and attempt to measure norms. It has become fairly common practice to refer to "social norms" or "internalized social norms" when discussing experimental (or field) data that appear unexplainable by existing economic theories of behavior.[5] Essentially, much of this literature has used "social norms" as a black box meant to capture some of the influence of the social environment on individuals' decisions. Only a few attempts have been made to obtain direct evidence about what the norm involves and whether the decision-makers involved perceive the norm to be at work.[6] We are perhaps the first to be able to differentiate clearly and in a conceptually-clean manner between social and moral norms in the laboratory. In this respect, a second major contribution of our paper is that our design allows us to observe both social and moral norms in the laboratory and to disentangle their effects.[7] Finally, while there is existing evidence that social observability induces less selfish behavior, we show that advice anchors the direction in which social observability affects behavior.

Ours is an attempt to open the black box and to provide experimental tools that allow researchers to better understand the content and consequences of social norms and to distinguish these from moral norms. In our experiments we use two tools to disentangle social and moral norms. First, we attempt to create a shared understanding about the action that is deemed to be appropriate. We do so by providing impartial advice by the decision-maker's peers. Second, we manipulate the social salience of a decision by varying the extent to which others can observe it. Note that this sets up a third characteristic of our design: because we know what advice is given, we can measure the shared understanding concerning appropriate action. Measuring this shared

---

[5] Combining "laboratory experiments", "economics", and "social norm" on Google Scholar in May 2011 yielded 541 academic papers since 2005.

[6] Some attempts that have been made will be discussed below.

[7] There are also other norm variants. For example, observing behavior of others (as in the case of conditional cooperation) can trigger quasi-moral norms; legal norms indicate explicit punishments for violations, while conventions require no sanctioning of violators. We exclude such norms by design.

understanding is a necessary condition for knowing the norm and, hence, for a proper analysis of its effects. Of course, whether impartial peer advice actually creates a shared understanding in this respect is an empirical matter (which we will address below). Given the results in the previous literature, we hypothesized that norms would work against purely selfish behavior.

One might therefore expect that choices are less selfish, both when advice is given and when choices are public rather than private; the effect might be largest when these forces are combined. We do in fact find a very significant effect of advice (i.e., the shared understanding), but only when the action taken is made public. This means that we do not observe moral norms *per se*, which would affect decisions even without observability. The combined effect is substantial, however, with considerably less self-interested behavior when there is both a shared understanding and behavior is observable. In addition, the results indicate that women are somewhat more sensitive to the advice received. These results are not consistent with the leading social-preference models, since these prescribe that one's own preferences should be unaffected by the desires of others who can observe the choice made, but who are not financially affected by it.[8]

The remainder of the paper is organized as follows. Section 2 presents the experimental literature on social norms, followed in section 3 by a simple model of preferences that allows for moral and social norms that affect choices. A presentation of our experimental design is given in section 4. We describe our experimental results in section 5 and offer some discussion and conclude in section 6.

---

[8] They may, however, be affected by one's *expectations* about another party's expectations, as in Charness and Dufwenberg (2006).

2. Previous Literature

Although the study of (social) norms and their effects has been much more common in sociology,[9] economists have increasingly acknowledged the important role they can play in economic decision-making.[10] Early attempts to theoretically capture this role include Akerlof (1980) and Lindbeck, Nyberg and Weibull (1999). Both show how social norms can affect behavior in the labor market in a way that contradicts traditional economic intuition. Akerlof (referring to "social customs") sees such norms as an explanation for the persistence of fair (as opposed to market-clearing) wages. Lindbeck *et al*. argue that social norms such as that one should "live off one's own work" help explain why modern welfare states need not collapse under the weight of excessive welfare claims.

Social norms and their enforcement are also often cited in work in experimental economics. For example, Ostrom (2000) argues that laboratory data show how norms can evolve to support cooperative behavior in common pool dilemmas. The literature on punishment in public goods games regenerated by Fehr and Gächter (2000a, 2000b) is based on the premise that punishment of low contributors serves to enforce a norm of social behavior.[11] Some noteworthy attempts in the experimental-economics literature to carefully define social norms implicitly make Elster's (2009) distinction between moral and social norms by stressing the 'social' part. Ostrom (2000) stresses: "social norms are *shared* understandings about actions that are obligatory, permitted, or forbidden" (pp. 143-144; italics are ours). Fehr and Gächter's (2000b)

---

[9] For an example and references, see Coleman (1988), which is a classic and widely-cited study that systematically analyzes social norms and their role in social capital. For a classic example in psychology, see Schwartz (1973).

[10] Bendor and Swistak (2001) investigate the issue of why there are norms, and show how social norms can be derived from principles of boundedly-rational choice as mechanisms that are needed to stabilize behaviors in many evolutionary games.

[11] An example of this line of literature is Fehr and Fischbacher (2004). They provide a major contribution to understanding the role of enforcement mechanisms in maintaining social norms, but rely on the "idea that social norms apply" (p. 65). Though such studies on enforcement are crucial for understanding why norms persist, they do not allow us to study the characteristics of norms *per se*. While the notion of social norms is frequently mentioned in the paper, the main analysis is not informative about the social norms themselves.

definition includes the notion that social norms involve a "shared belief" and are enforced by "informal social sanctions". Finally, in their classic study on giving in dictator games, Hoffman, McCabe and Smith (1996) manipulate the effects of (social) norms in a design that varies the social distance between the dictator and the experimenter. This implicitly focuses on the 'social' aspect of these norms.

As mentioned above, norms remain a black box in much of the literature. More specifically, we distinguish three problems. First of all, it often happens that the introduction of a norm as an explanation for observed behavior is made without a discussion of (i) why a norm might develop; (ii) what the norm involves; and (iii) why it should apply to the experimental environment under investigation. Second, authors rarely define what they mean by a (social or other) norm; notable exceptions are discussed below. The relationship between a norm and behavior then remains at an intuitive level. Third, because the experimenter does not have direct information about the norm that an individual decision-maker feels is applicable to an experimental environment, this norm then becomes a 'homegrown restriction' on the values induced in the experiment. This implies a loss of experimental control and may therefore harm the internal validity of the experimental design. As will become clear, our design avoids these three problems.

Whereas ours is an attempt to directly measure norms in a laboratory environment, the typical scenario in previous attempts has been to indirectly derive the norms that are at work from subjects' choices (cf. the Fischbacher and Fehr 2004 example in fn. 8). An exception to this practice is recent work by Erin Krupka and her co-authors (e.g., Krupka and Weber 2010; Burks and Krupka 2011). They attempt to derive independent information about norms by having a group of subjects participate in set of coordination games. The basic idea is as follows. They measure the extent to which actions are socially appropriate or inappropriate by presenting

6

respondents with a description of a choice environment, including a set of possible available actions. Then everyone is asked to evaluate the (in)appropriateness of this action. If one's appropriateness rating for a given action is the same as the modal response in the experiment, then one obtains a monetary prize. The authors assume that a social norm provides a focal point to make the choice most in line with this norm the expected equilibrium in the coordination game. Note, however, that this assumption cannot be directly corroborated. Coordination on any response category is in fact an equilibrium of this game. Moreover, information about the norm is dependent on the cases selected by the researcher in the first place.[12] Nevertheless, this method provides an interesting attempt to incentivize the elicitation of information on the content of social norms. Note that our elicitation of this information is much more direct, i.e., through peer advice. To discover the norm, we simply ask advisors to tell others what the norm is.

Reuben and Riedl (2011) is an example of a study that combines independent (direct) elicitation of norms with indirect derivation from subjects' choices. They investigate contribution norms in a laboratory public goods game and use a (non-incentivized) survey of individuals from the same subject pool to collect independent information about the norms that apply. They use their laboratory data on punishment to determine which of the norms elicited in the survey play a role. Their definition of a social norm is adjusted to this interactive setting. They define a social norm as a behavioral rule that is, first, known to exist; and second, motivates those involved to follow the rule under the condition that (a) it is believed that sufficiently many others also do so, and either (b) it is believed that sufficiently many others expect one to follow the rule, or (b') it is believed that sufficiently many others are willing to sanction deviations from the rule.[13] Note

---

[12] Krupka and Weber (2010) are able to give the entire action space for rating. This is not the case in the other papers cited, however, and seems to be an exceptional case.

[13] This definition allows for the possibility that a norm exists without being adhered to. Some argue that one should distinguish between 'descriptive norms' and 'injunctive norms' (e.g., Bicchieri 2006, Cialdini et al. 1990, Deutsch and Gerard 1955).. The latter prescribe behavior, while descriptive norms basically describe what people do. It is not clear to us that this distinction is

that it remains unclear in this setup whether the punisher and person being punished share an understanding about what one ought to do. For example, a participant in these experiments may well disagree that a contribution norm exists, but adhere to it to avoid the costs of being punished.

The literature also shows studies where either of our tools (advice or observed choices) is used in a different context. For example, there have been previous experiments involving an agent (or agents) offering advice or expressing preferences. These have involved self-interested agents, however, whose material payoffs depend on the choice made by the person receiving advice.[14] However, it is not obvious that a decision-maker would or should consider advice by an interested party to constitute a shared understanding about what one ought to do in a specific situation.[15] There are also experiments involving a third (or even fourth) party who observes (but is not directly affected by) the choice by an agent and can choose to punish perceived malfeasance; note that these do not include a shared understanding. These studies show that impartial observers are willing to punish, presumably when they feel that a norm has been violated.[16] They do not, however, create a shared understanding about what the norm is, nor do they provide the experimenter with information about this norm.

useful. In Elster's definition all norms are injunctive. This does not mean that people adhere to them in all situations, however. For example, in a standard model of preferences and restrictions, selfish behavior may be chosen even if one acknowledges that the norm is to act unselfishly. An example of such a model is presented in section 3.

[14] Schotter and Sopher (2003, 2006, 2007) and Chaudhuri, Schotter, and Sopher (2009) investigate the effect of advice in 'intergenerational' games in which the advisor has a financial interest in the strategic decision made by the advisee. In Charness and Rabin (2005), recipients in dictator games and responders in sequential games can express their preferences over the binary choices available to the other player.

[15] Konow (2000, 2003) provides extensive discussions about distributive justice. One related point is that being overly self-interested can result in a form of cognitive dissonance. One would expect this dissonance to be greater when norms are made more salient and the action taken is nevertheless self-interested.

[16] Observer (or audience) effects that did not involve direct punishment have been observed in Charness, Rigotti, and Rustichini (2007) and Andreoni and Bernheim (2009). Fehr and Fischbacher (2004) allow third-party observers of a Prisoner's dilemma to punish (at a cost) either player, with about half punishing a defector when the other player has cooperated. Charness, Cobo-Reyes, and Jiménez (2008) investigate costly third-party punishment and reward in a 'trust' game, finding that responders return a higher proportion of the amount sent to them when there is the possibility of punishment. In Coffman (2009), a fourth party can punish (at no cost) a dictator who has either chosen an allocation or chosen to unilaterally 'sell' the game to a second party; in either case, a third person receives the eventual allocation made. There is less punishment when the dictator delegates.

So while there are experiments in which self-interested parties make recommendations and experiments in which non-self-interested observers can choose monetary punishments, we are unaware of any study in which a shared understanding is induced through public and impartial advice and people are subject to only social or moral sanctions.[17]

3. A Model of Social and Moral Norms

We first introduce the notion that an individual (i) may care about her (self or public) image, ξ_i. When making a choice $c_i \in C$, she derives utility from the own earnings that follow from c_i, but also from a positive image that may result from her choice. This gives:

$$\boxed{} \tag{1}$$

where U denotes utility, $\pi(\bullet)$ is i's monetary payoff, $\gamma_i \geq 0$ measures her sensitivity to image, p is a dummy indicating whether her decision is observable by others ($p=1$) or not ($p=0$), and π_n is her payoff if she follows the existing norm with respect to choosing from C. When making the decision, i's image has not yet been formed, hence her expected image enters her utility in (1).

We assume that one's (expected) image ξ_i is positively correlated with the closeness of the payoff related to choice c_i to the norm π_n,[18] i.e.,

$$\frac{\partial E(\xi_i)}{\partial \left(\left| \pi(c_i) - \pi_n \right| \right)} < 0. \tag{2}$$

Furthermore, with public observation of decisions ($p=1$), both expected self-image and expected public image play a role, whereas only self-image can play a role with private decisions. Hence:

$$\gamma_i(1) > \gamma_i(0) \tag{3}$$

---

[17] The closest is a study by Masclet, Noussair, Tucker, and Villeval (2003) that involve non-monetary sanctions (costless '"punishment points") in a public-goods game; these sanctions are related to violations of social norms. However, no shared understandings are induced.

[18] Note that we simplify matters by assuming that image is based on the own payoff related to i's choice as opposed to the choice itself. Moreover, (2) assumes that any deviation from the norm negatively affects image. This creates the possibility for anti-social punishment as observed by Hermann et al. (2008).

A simple quadratic specification of (1) assumes that the expected image varies negatively with the squared difference between chosen own payoff and the norm. Utility is then given by:

$$\boxed{}.\tag{4}$$

Note that we have normalized the expected image to zero if i behaves according to the norm. If she deviates from the norm, this reduces her utility when $\gamma_i > 0$. The marginal disutility increases with the deviation.

To implement (4), one needs to specify the norm payoff, π_n. Recall that the norm is characterized by a shared understanding of what one ought to do. We distinguish between the case where the decision maker receives a direct signal about what her peers believe and the case where she does not receive such a signal. The direct signal is represented by an option $\alpha_n \in C$. When she does receive such a signal, this constitutes the norm for her:

$$\pi_n = \pi(\alpha_n).\tag{5a}$$

Without direct signal, i must form an expectation about what others think she ought to do. We assume she has accurate beliefs about what her peers believe and estimate this by the expected signal for the situation she is in ($p = 0$ or $p = 1$):

$$\pi_n = \sum_{a \in C} f_p(a)\pi(a),\tag{5b}$$

where $f_p(a)$ is the relative frequency with which her peers believe a is what one ought to choose.

The challenge for an empirical application is to have a measure for either α_i or $f_p(a)$. Our experimental design provides us with such measures. After introducing this design, we will apply this model to our setting.

4. Experimental Design Issues and Implementation

Our experiments were conducted at the University of Amsterdam, using a 2x2 design that varied whether advice was given and whether the dictator action taken was made public. There were 345 participants in total; 255 of these were in choice groups, while 90 only gave advice. Each participant received an initial endowment of €16. The experiment took approximately 25 minutes and the average earnings were €14.66.

People were formed into groups of three. There were 16 sessions in all. Table 1 summarizes the numbers of observations per treatment. Each 'advice group' gave advice to one or two other groups (depending on the number of participants in a session), with the advice necessarily the same in both cases when there were two. Specifically, with 3N participants, the groups were randomly numbered 1, 2, ….N. Group 1 gave advice to groups 2 and 3, group 4 gave advice to groups 5 and 6, etc. If there were 15 participants, group 4 only gave advice to group 5. If there were 21 participants, group 7 formulated a recommendation but this was not passed on to any other group. The members of an advice group could communicate via chat box for 60 seconds and then make a selection (if each person made a different choice, another 60 seconds of chat would follow; in practice, this was never necessary). The members of a choice group did not communicate; instead, one of the three members was randomly appointed to be a dictator and made a choice unilaterally.

Table 1: Treatments

	Private Payoff	Public Payoff
No Advice	**NPr** 17 dictators	**NPu** 20 dictators
Advice	**APr** 25 dictators, 17 advice groups	**APu** 23 dictators, 13 advice groups

*Because of subject turnout, some groups' advice was not passed on to any choice group (see the main text). Two advice groups were not used in APr and one advice group was not used in APu.

Each dictator made a choice from six possible alternatives, which varied in terms of own payoff consequences and the payoff consequences for the other two people in her choice group. These feasible choices are shown in Table 2.

Table 2: Choices available to the dictator

Choice	A	B	C	D	E	F
Own payoff	8	4	0	-4	-8	-12
Payoff to player 2	0	4	8	8	4	0
Payoff to player 3	-16	-12	-8	-4	0	4

First, we briefly mention what various models of distributional preferences indicate would be chosen by a non-self-interested party (like a member of a choice group).[19]   A (hypothetical) non-self-interested central planner with Fehr and Schmidt (1999) preferences would minimize the sum of the pairwise differences between payoffs, so that option D or E is best.[20] If this non-self-interested person had Charness and Rabin (2002) quasi-maximin (social-welfare) preferences, both the sum of the payoffs and the minimum payoff would come into play. Both C and D provide a total payoff of 0, higher than any of the other choices; D also provides the highest minimum payoff (-4) of any of the choices. Thus, D should be chosen with Charness-Rabin distributional social preferences.

Predictions for the dictators in the choice groups will vary according to the parameter values of the social preferences under consideration. For example, the choice by a dictator with

---

[19] Previous work on this issue includes Frohlich and Oppenheimer (1984, 1992), who examine the allocations made to a group of individuals by decision-makers who do not know which payoff would be theirs, and Charness and Rabin (2002), who consider allocations made by an individual who does not know his or her own payoff, but chooses for two anonymous others (see games Barc10 and Barc12). These studies find that people reach agreements that tend to maximize total payoffs, while observing an income floor for individuals in the group.

[20] If this person instead had Bolton and Ockenfels (2000) preferences, he or she would consider the ratio of one person's payoff to the total, seeking to keep these as close to 1/3 as possible. However, this is less easy to analyze. For example, the total payoff is the case of options C or D is 0, so that ratios are somewhat arbitrary. Thus, we do not discuss this model further.

standard Fehr-Schmidt preferences depends on the extent of inequity aversion. With an advantageous inequity-aversion parameter equal to 0.6 and advantageous inequity parameter equal to 1 or 4 (40% of the subjects categorized in Fehr and Schmidt 1999) would choose B. Lower inequity aversion would lead to A being chosen.

All in all, non-self-interested advice will tend to allocate less to the dictator than she would choose herself, even if she has pro-social preferences. If this advice creates a shared understanding, then the corresponding social norm will tend to make the dictator's choice less selfish (i.e., a column more to the right in Table 2).

The experimental procedures were as follows. After being randomly seated, players received computerized instructions (see Appendix A for a translation). These stated that they would be allocated to groups of three participants (each group consisting of a Red, a Blue, and a Green player). They were informed that Blue and Green would not make any decision in the experiment and Red would decide on a column of the payoff matrix given above (Table 2). They knew that this was a one-shot decision only. In the treatments with advice, each person also received these instructions:

"We imagine that it may be difficult for Red [the dictator] to decide what one ought to do in this situation. We will therefore randomly select a few groups and have them formulate advice. The groups that give advice will not subsequently choose from the options. Each member of an advice group will therefore earn exactly €16 in this experiment. The advice by every group may be passed on to more than one group. The advice reads: 'The group that was asked to advise informs you that red ought to choose … in this situation [where we substitute the advised option (A, B, C, D, E, F) for the dots]'".

13

In the case of public payoffs, further instructions read:

> "Today's payoffs are organized differently than you may be used to at CREED. At the end of the experiment, you will be called forward one by one. We will then publicly announce your role today, what you decided, and how much you earned. Payoffs today are therefore not private and anonymous."

When a participant came forward, all monitors showed the role and decision. Participants had to return to their seats and wait until everyone had been paid.

In terms of what conclusions can be drawn from comparative statics, we have four direct comparisons. Going from NPr to APr reflects induced moral norms, going from NPu to APu reflects induced moral norms or induced social norms, going from NPr to NPu reflects homegrown social norms, and going from Apr to APu reflects homegrown or induced social norms (holding the advice constant).[21] For example, comparing differences between NPr and APr and between NPu and APu will allow us to isolate the additional effect of social norms from the effect of moral norms on behavior.[22]

For more specific predictions, we use the model presented in the previous section. With (4) and (5a) we can determine the utility obtained from choosing the distinct options for the treatments with advice. In doing so, we assume that the advice constitutes the signal α_n in (5a). The utility then depends both on the option and on the advice received. For each combination, the utility is shown in Table 3.

---

[21] By 'homegrown' we mean a norm concerning this situation that an individual brought into the laboratory. By 'induced' we mean a norm that was created in the laboratory.
[22] Our design will not allow us to distinguish between induced moral norms and some other effect of (private) advice. It is difficult to imagine ways to make this distinction, however.

Table 3: Utility derived from choices

Advice: Choice:	A	B	C	D	E	F
A	1	$1-0.25\gamma_i$	$1-\gamma_i$	$1-2.25\gamma_i$	$1-4\gamma_i$	$1-6.25\gamma_i$
B	$0.5-0.25\gamma_i$	0.5	$0.5-0.25\gamma_i$	$0.5-\gamma_i$	$0.5-2.25\gamma_i$	$0.5-4\gamma_i$
C	$-\gamma_i$	$-0.25\gamma_i$	0	$-0.25\gamma_i$	$-\gamma_i$	$-2.25\gamma_i$
D	$-0.5-2.25\gamma_i$	$-0.5-\gamma_i$	$-0.5-0.25\gamma_i$	-0.5	$-0.5-0.25\gamma_i$	$-0.5-\gamma_i$
E	$-1-4\gamma_i$	$-1-2.25\gamma_i$	$-1-\gamma_i$	$-1-0.25\gamma_i$	-1	$-1-0.25\gamma_i$
F	$-1.5-6.25\gamma_i$	$-1.5-4\gamma_i$	$-1.5-2.25$	$-1.5-\gamma_i$	$-1.5-0.25\gamma_i$	-1.5

Notes: Without loss of generality, we have scaled payoff by dividing them by 8, which normalizes the maximum private payoff to 1.

One first notes that the model predicts that no one will ever choose less selfishly than advised. For any γ_i, and any choice deviating from the advice by giving up more of the own payoff than advised, a higher payoff can be obtained by deviating the same distance towards a more a selfish choice. Table 3 allows us to determine γ_i from the choice made. Table 4 shows the interval implied by the choice.

Table 4: Estimates for γ_i, advice treatments

Advice: Choice:	A	B	C	D	E	F
A	$[0,\infty)$	$[0,2)$	$[0,0.67)$	$[0,0.4)$	$[0,0.29)$	$[0,0.22)$
B	––	$[2,\infty)$	$[0.67,2)$	$[0.4,0.67)$	$[0.29,0.4)$	$[0.22,0.29)$
C	––	––	$[2,\infty)$	$[0.67,2)$	$[0.4,0.67)$	$[0.29,0.4)$
D	––	––	––	$[2,\infty)$	$[0.67,2)$	$[0.4,0.67)$
E	––	––	––	––	$[2,\infty)$	$[0.67,2)$
F	––	––	––	––	––	$[2,\infty)$

Notes: The Table gives γ-values for normalized payoffs. To obtain the corresponding estimates for unscaled payoffs, multiply these numbers by 64.

A similar exercise can be done for the treatment without advice. For this purpose, we determine π_n using (5b) and the frequency with which we observe various advice (which will be presented in the results section, see table 7. For normalized payoffs, this gives:

$$\pi_n\big|_{p=1} \approx 0.04$$
$$\pi_n\big|_{p=0} \approx 0.33$$

Table 5 then gives the utility and γ-estimates for distinct choices in the two treatments.

Table 5: Estimates for γ_i, no-advice treatments

	Public Payoff		Private Payoff	
Choice	Utility	γ_i	Utility	γ_i
A	$1-0.92\gamma_i$	$[0,0.70)$	$1-0.45\gamma_i$	$[0,1.14)$
B	$0.5-0.21\gamma_i$	$[0.70,2.27)$	$0.5-0.01\gamma_i$	$[1.14, \infty)$
C	$-0.02\gamma_i$	$[2.27, \infty)$	$-0.11\gamma_i$	---
D	$-0.5-0.29\gamma_i$	---	$-0.5-0.69\gamma_i$	---
E	$-1-1.08\gamma_i$	---	$-1-1.77\gamma_i$	---
F	$-1.5-2.37\gamma_i$	---	$-1.5-3.35\gamma_i$	---

Notes: For each payoff treatment, the first column gives the utility attributed to a choice and the second column shows the set of γ-values for which this choice is optimal. Note that the model predicts that options D-F will not be chosen for any γ.

Comparing the values for γ_i with public and private payoffs, it is clear that the distribution of choices with public payoffs stochastically dominates the distribution with private payoffs, which shows that public image matters much more than private image. For values between 0.70 and 1.14, a person would choose B with public payoffs, but A with private payoffs. Similarly, with values above 2.27, a person would choose C with public payoffs, but B with private payoffs.

5. Results

We find substantial differences in the allocation choice selected, depending on whether advice was given and whether payoffs were public or private. Table 6 presents the percentage of choices in each category for each treatment.

Table 6: Distribution of choices, by treatment

	A	B	C	D
No advice, private	.706	.294	.000	.000
No advice, public	.550	.350	.100	.000
Advice, private	.640	.320	.040	.000
Advice, public	.391	.217	.261	.131

We see that no choice of E or F was ever made. Furthermore, the only instances in which a dictator chose a negative own payoff (option D) was when advice was given and the choice was public. When there was no advice and payoffs were private, more than 70% of all choices were entirely selfish, with the remainder in the next-most-selfish category. Across the four treatments, the differences are significantly different (Kruskal Wallis, χ^2=9.179, two-tailed p=0.027.[23]

Pairwise Wilcoxon ranksum tests across treatments show significant differences between the choices made in the advice/public treatment and each of the other three treatments. We find $Z = 2.513$, $p = 0.006$, for advice/public versus no advice/private, $Z = 1.722$, $p = 0.043$, for advice/public versus no advice/public, and $Z = 2.404$, $p = 0.008$, for advice/public versus advice/private (all tests are one-tailed, in keeping with our directional hypotheses). None of the other pairwise comparisons are significant ($Z = 1.127$, $Z = 0.515$, and $Z = -0.714$ for the comparisons between no advice/public and no advice/private, advice/private and no advice/private, and advice/private and no advice/public, respectively).

If we compare all choices with public payoffs to all choices with private payoffs, the ranksum test gives $Z = 2.457$, $p = 0.007$ (one-tailed test). If we compare all choices with advice to all choices with no advice, the ranksum test gives $Z = 1.388$, $p = 0.083$ (one-tailed test). So at first glance it seems that observability has a stronger effect than advice. Nevertheless, advice is

---

[23] For numerical analysis we convert the options to numeric form such that A=1, B=2, etc.

important and it is instructive to see its effects and to test whether the advice given is independent of whether payoffs are public or private.

In fact, taking the last question first, advice is quite sensitive to the exposure of the choices made.[24] When the payoffs were private, two groups recommended A, seven groups recommended B, four groups recommended C, and one group recommended D; when the payoffs were public, two groups recommended A, one group recommended B, seven groups recommended C, and three groups recommended D. A Chi-square test (grouping A with B and C with D, so as to have a sufficient number of observations in each cell) gives $\chi_1^2 = 4.638$, $p = 0.031$, so that advice is significantly more in the direction of non-selfishness when the payoffs are public. Thus, perhaps surprisingly (since the advice is not made public), the advice given depends on whether the subsequent choice will be made public.[25] One possible explanation is that advisors themselves distinguish between moral and social norms. We will discuss this in more detail, below.[26]

How do choosers react to advice? Table 7 shows the responses to specific advice in the treatments with public and private payoffs. Eighteen of 23 groups (78.3%) received recommendations (from 13 distinct advice groups) to choose C or D with public payoffs, compared to eight of 25 groups (32.0%) receiving recommendations (from 14 advice groups) with private payoffs.[27] This difference in proportions is statistically significantly ($Z = 3.213$, $p =$

---

[24] In the following analysis we disregard advice groups whose advice was not passed on to any choice group (see table 1).

[25] Interestingly, the degree of consensus on the norm also appears to be less clear in the public case (note that those receiving advice did not know whether it was unanimous). There were 12 advice groups with private payoff. Three of these had a split decision (2-1). There were 17 advice groups with public payoff and 8 of these had a split decision. The difference in rates, while large (47% versus 25%), is not statistically significant ($Z = 1.21$). This strengthens our results, because the public case is where we found that a norm has an effect.

[26] The 'selfish' option A is advised twice in each treatment. The main difference is that B is advised seven times when decisions are private and only once for the public-payoff treatment. C and D are advised more often with public payoffs (seven vs. four and three vs. one, respectively).

[27] Note that the numbers concerning advice received are different than the numbers given in the previous paragraph on advice given because advice was often given to two groups.

Advice given, Public payoffs	Number of responses			
	A	B	C	D
A	2		1	
B	1	1		
C	4	4	4	
D	2		1	3

Advice given, Private payoffs	Number of responses			
	A	B	C	D
A	3		1	
B	7	6		
C	5	1		
D	1		1	

0.001, one-tailed test), a matter that we must take into account when interpreting advice as inducing norms.

Many people (10 of 23, or 43.5%, with public payoffs; nine of 25, or 36.0%, with Private payoffs) respond to the advice by choosing it.[28] Of the remaining choice groups, in only one case in each treatment does any group choose to be less self-interested, with the remaining groups choosing to be more self-interested than the recommendation; the binomial test finds that this deviation pattern far from random ($Z = 5.014$, $p < 0.001$). The difference between the advice given and the choice is 0.783 categories on average (standard error $= 0.243$) with public payoffs. With private payoffs the average difference is 0.840 categories (standard error $= 0.189$). Thus, there is no difference across treatments in the deviation from the advice given.

Regression analysis

We supplement the non-parametric analysis shown above with formal regression analysis, allowing us to account for effects from different variables in a comprehensive framework. Table

---

[28] To the extent that stakeholders such as these dictators are self-deceptive (see for example Konow, 2000), having an impartial third party increases the 'cost' of self-deception, resulting in less selfish behavior.

8 considers the determinants of the choice made, as well as the determinants of how the choice made differs from the advice received. In these regressions, we again convert the choices into numeric form (A = 1, … ,F = 6). Our rationale for using an ordered-probit regression model is that categories represent increasing amounts given up by the dictator. The baseline is private payoffs with no advice.

Table 8: Ordered-probit regressions for choice and for choice-advice differences

Dependent variable	Chosen option (1)	Chosen option (2)	Chosen option (3)	Deviation from advice (4)
Sample	No advice received	Advice received	Advice received	Advice received
Public (dummy)	0.534 (0.413)	1.129 (0.373)***	0.924 (0.393)**	0.944 (0.378)**
Advice received	---	---	0.336 (0.202)*	−1.106 (0.245)***
Female	−0.239 (0.410)	0.643 (0.361)*	0.619 (0.362)*	0.604 (0.345)*
#obs.	37	48	48	48

Notes: ***, **,* denote significance at p = 0.01, 0.05, and 0.10, respectively (two-tailed tests). Independent variables are: *Public* = dummy variable equal to one in the public payoff treatment; *Advice (dummy)* = dummy variable equal to one in the treatments where advice was given; *Advice received* = the numerical value of the advice received; *female* = dummy variable equal to one if the dictator is female.

Specifications (1) and (2) consider the effect of observability on the dictator's decision. Specification (1) does so for the subsample that did not receive advice and specification (2) for the dictators with advice from their peers. The results clearly show that dictators make less self-interested choices when their decisions will be made public, but only if they have received advice on what one ought to do; it also indicates that women tend to choose higher numbers than men when their decisions are observed (although the coefficient is only marginally significant). In specification (3) we add the content of the advice as an explanatory variable. The results again confirm that dictators choose less in line with their self-interest when the advice is public and show that their choices correlate (marginally significantly) positively with the advice given. Again, women tend to choose less selfishly than men. Considering the extent to which the choice deviates from the advice received, model (4) uses Advice minus Choice as the dependent variable and shows that the more that the advice suggests one sacrifice, the more that dictators

deviate from this advice; moreover, observability decreases this deviation and females are more sensitive to advice and deviate less from it than men. All in all, having public payoff strongly affects dictators' choices by moving them in the direction of the shared understanding created by advice. Without advice, making the payoff public has no effect. Finally, observe that the differences in advice between the public and private treatments are not driving the observed differences in dictator's choices. Specification (4) shows that deviations from advice are lower when choices are observed.

To interpret our results in terms of moral versus social norms, we return to our design table and enter the average numerical values of dictators' choices (Table 9).

Table 9: Average Choices

	Private Payoff	Public Payoff
No Advice	**NPr**: 1.3	**NPu**: 1.6
Advice	**APr**: 1.4	**APu**: 2.1

The number in each cell is the average numerical value of the choice made.

As discussed, the difference between the condition with advice with public payoffs (APu) and each and every other condition is statistically significant, while none of the other pairwise comparisons are significantly different from zero.[29] As argued above, we can conclude from the significant change when we move from NPu to APu that there is either an induced moral norm or an induced social norm. From the significant effect from APr to APu, we conclude that (holding advice constant) there are either homegrown social norms at work, or we have induced a social norm. On the other hand, the lack of difference between NPr and APr implies that there are no induced moral norms/effect of advice and the lack of difference between NPr and NPu means

---

[29] Note that the difference between the average choices between the treatments with no advice and advice is much larger (0.5 versus 0.1) with public payoffs than with private payoffs, further illustrating that advice matters primarily when payoffs are public.

that there is no homegrown social norm at work.[30] The explanation that still stands is that the combination of advice (shared understanding) and public payoff (observability) has induced a social norm in the laboratory to which participants adhere.

To further investigate the extent to which we have created a social norm in the laboratory, we need to investigate whether we have created both observability and a shared understanding about what one ought to do. The former seems trivial in our public-payoff treatment. As for a shared understanding, first note that we purposefully used a game where the social norm is *a priori* unclear, hence a shared norm does not seem to exist (though individual subjects may believe that it does). This is apparent, for example, from the large diversity in advice given. To investigate whether we created a shared understanding, we use responses to a post-experimental questionnaire. One of the questions asked was: "What do you think red should choose in this experiment?". This question does not directly ask about a perceived norm but combines the effect of a (social) norm governing red's choice with any perceived tradeoff for red between this norm and selfish considerations (as in eq. 4, for example). If the advice creates a shared understanding about the norm, then a response to this question should be positively related to the advice received. All participants responded to this. Regressing the answer given on the advice received (using ordered probit) yields a positive coefficient of 0.400 with $p < 0.001$. Hence, the advice appears to strongly affect what participants believe red should do.

For a more detailed look at this effect of advice, we dichotomize the decisions (and advice) into 'selfish' (categories A or B) and 'unselfish' (C or D). We then consider whether advice to be unselfish will lead to more responses that red should choose unselfishly. Figure 1

---

[30] It is also possible that advice establishes moral norms, but that these have no effect; however, this goes against considerable evidence in the social sciences. Further, given the effect of induced social norms, we suspect that it is the case that advice does not establish moral norms.

show the results, disaggregated by the observability of decisions and the role of the respondent in the experiment.

Figure 1: Effect of Advice on Shared Understanding
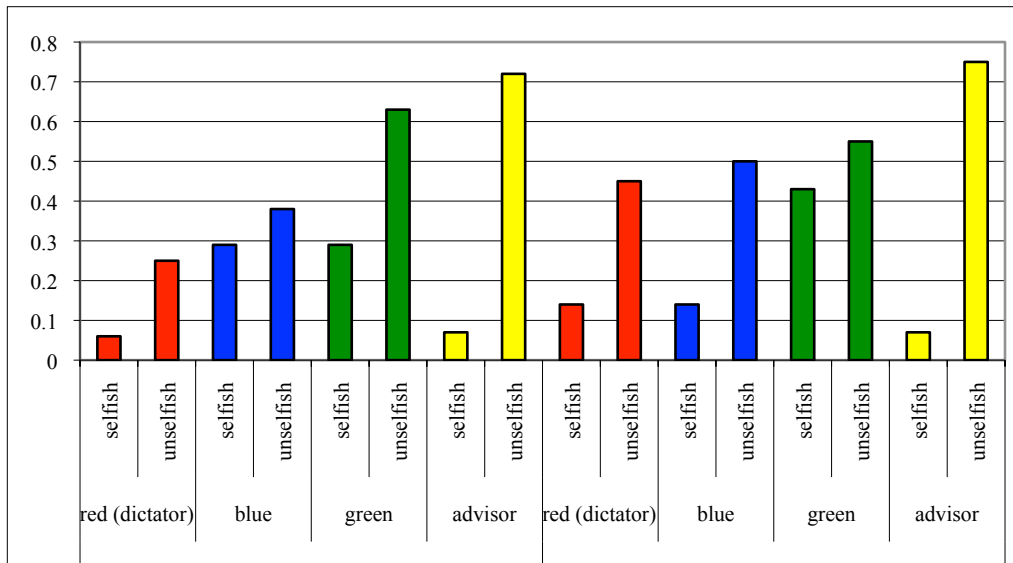
Notes: bars show the fraction of respondents that responds with categories C or D to a post-experimental question on what the right thing is for red to do. "Selfish" ("unselfish") indicates that the advice was to choose A or B (C or D). A color indicates the respondent's role in the group, where an 'advisor' was member of an advice-giving group. The distinction private-public refers to the unobservability vs. observability of choices.

Note from the figure that after receiving unselfish advice respondents believe that Red should choose less selfishly than after selfish advice. This holds in all cases. The difference is larger when decisions are observed than when they are not (with the exception of Green participants). The difference is largest for advisors. This is not surprising, since they are the ones having given the advice in the first place.

All in all, from this post-experimental questionnaire it appears that the advice given does coordinate beliefs about what the dictator ought to do. In this respect it contributes to a shared

understanding and, hence, a norm. This effect is strongest when the norm is social.[31] A violation of this norm means a violation of expectations and so could be related to simple guilt aversion (Charness and Dufwenberg, 2006), since the beliefs of another party (or parties) becomes part of one's own utility function and preferences; endogenous social norms are also part of the full Charness and Rabin (2002) model (with respect to "demerits"). However, neither of these models (nor the purely consequential social-preference models) picks up the effect of observability.

Finally, we can use observed choices to estimate image concerns (γ-values) for the utility function depicted in (4). Using the midpoints in the intervals defined in tables 4 and 5 and the frequencies reported in Tables 6 and 7, we determine the estimated distributions of the sensitivity to image, $\gamma(0)$ (private payoff) and $\gamma(1)$ (public payoff). These are shown in figure 2.

Figure 2: distribution of *image sensitivity*
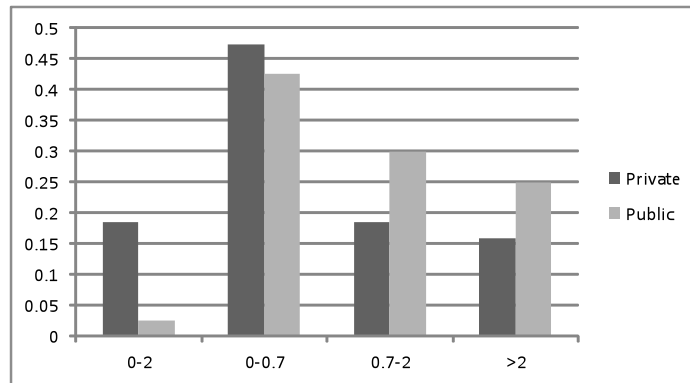
Notes: Bars show fractions of subjects with normalized γ-values in the ranges
indicated on the horizontal axis. Subjects that received advice A are not included. The
5 subjects in the NPr treatment with estimated $\gamma > 1.14$ are categorized as $\gamma \in [0.7,2]$
and the 12 subjects with range $\gamma < 1.14$ are categorized as $\gamma \in [0,0.7]$.

---

[31] Of course, one should take care in interpreting responses to a hypothetical question. Nevertheless, the correlation between advice and stated understandings is remarkable. Moreover, it is supported by the fact that actual dictator choices responded to advice in a similar way.

The figure clearly shows that the sensitivity to image is stronger for public image than for private image. For example, whereas 25% of the subjects with public payoffs have a normalized γ that is larger than 2, this holds for only 15% of the subjects with private payoff. Apparently, public image matters more to our subjects than self-image. Note that the difference in image sensitivity between public and private payoffs is statistically significant. This is because these estimates are based on observed choices and we reported above that the choices with public payoff are significantly less selfish than the choices with private payoffs ($Z = 2.457$, $p = 0.007$).

Recall that γ_i is defined in (3) as the coefficient on the squared difference between the chosen payoff and the normative payoff. Thus, with normalized payoffs (i.e, the payoffs from table 2 divided by 8), a dictator who chooses B when the norm indicates C would gain 0.5 payoff unit, but would lose $0.25*\gamma_i$ units.[32] So any dictator with $\gamma_i < 1/3$ should choose A, anyone with $\gamma_i \in [1/3, 2]$ should choose B and dictators with $\gamma_i > 2$ should choose C.

6. Conclusion

The influence of norms on behavior is an important issue in economics, with applications to areas such as labor markets, welfare economics, and common pool dilemmas. Elster (2009) defines a social norm as an injunction on behavior that is sustained by the threat of sanctions or social disapproval. He defines moral norms as internalized norms that are sustained by guilt or remorse and do not depend on observation. Since agents' behavior in the field may or may not be observed, both types of norm can potentially come into play in field settings.

---

[32] Using eq. (4): $U_A = 8/8 - \gamma_i*(0-8/8)^2 = 1-\gamma_i$; $U_B = 4/8 - \gamma_i*(0-4/8)^2 = 0.5-0.25\,\gamma_i$; $U_C = 0$.

We use a novel experimental design to study the influence of norms on behavior and to distinguish between social and moral norms. We have argued that a social norm requires a shared understanding about what one ought to do in a specific situation as well as observability of one's actions. In our design, there are degrees to which a decision-maker can sacrifice own payoffs to achieve an outcome that is more favorable to the other paired participants. A priori, there appears to be no clear norm to guide one's choice; here we induce a norm by providing impartial advice from one's peers. We have provided evidence that this advice creates the shared understandings needed to create a norm. This is made a social norm when one knows that one's choice will face public inspection. Indeed, we find that when one's decision is observed, there is a tendency to choose more in line with the social norm induced by advice. A moral norm is induced through advice given to the dictator if her choice is not made public. We observe no such effect of advice without observation.

Previous work has generally focused on social norms, without making explicit what are these norms or why they would play a role if others do not observe behavior. Studies involving advice have always involved self-interest on the part of the advisor, so that it remains unclear whether they create a shared understanding. Our design permits us to identify and disentangle the separate effects of social and moral norms by carefully creating a shared understanding and manipulating the observability of decisions; to the best of our knowledge, our study is the first to do so.

Social norms are found to affect behavior. While many previous studies have provided indirect evidence for this, our design allows us to show a direct link from the creation of a norm and the observability of choices to its effect on behavior. The combination of advice and public observation is particularly strong, in part because the advice (the norm that is created) is to be

26

more self-sacrificing when the advisor group knows that the choice will be made public; this implies that the advisor group, while anonymous, may well feel more responsibility for the outcome in this case, despite the fact that the advice given is not made public.

The model we have presented offers a simple structure to study the effects of norms in a simple economic framework. It attributes the key effects of norms to the utility attributed to fulfilling either the public image or the self-image that follow from one's actions. In the model, both factors may be important, allowing moral suasion to be complementary to self-reputational concerns.

While we have identified substantial effects from social norms in our data, it is clear that there is a great deal of research remaining in this important area. For one thing, the fact that moral norms do not play a role may depend on the decision under scrutiny. There may be environments where moral norms can be induced that affect individuals' choices. We do believe that only careful experimental manipulation like the one we have presented here can help discover these environments.

References

Akerlof, G. (1980), "A theory of social custom, of which unemployment may be one consequence," *Quarterly Journal of Economics*, **94(4)**, 749-775.

Andreoni, J. and D. Bernheim (2009), "Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects," *Econometrica*, **77(5)**, 1607-1636.

Bendor, J. and P. Swistak (2001), "The evolution of norms," *American Journal of Sociology*, **106(6)**, 1493-1545.

Bicchieri, C. (2006). *The Grammar of Society: The Nature and Dynamics of Social Norms*, New York: Cambridge University Press.

Binmore, K. and L. Samuelson (1994), "An Economist's Perspective on the Evolution of Norms," *Journal of Institutional and Theoretical Economics*, 150:45-63

Bolton, G. and A. Ockenfels (2000), "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Review*, **90(1)**, 166-193.

Burks, S. and E. Krupka (2011), "A Multi-method Approach to Identifying Norms and Normative Expectations within a Corporate Hierarchy: Evidence from the Financial Services Industry", paper presented at the 2011 THEEM conference, Kreuzlingen.

Charness, G., R. Cobo-Reyes, and N. Jiménez (2008), "An investment game with third-party intervention," *Journal of Economic Behavior and Organization*, **68(1)**, 18-28.

Charness, G. and M. Dufwenberg (2006), "Promises and Partnership," *Econometrica*, **74(6)**, 1579-1601.

Charness, G. and M. Jackson (2009), "The role of responsibility in strategic risk-taking," *Journal of Economic Behavior and Organization*, **69(3)**, 241-247.

Charness, G. and M. Rabin (2002), "Understanding social preferences with simple tests," *Quarterly Journal of Economics*, **117(3)**, 817-869.

Charness, G. and M. Rabin (2005), "Expressed preferences and behavior in experimental games," *Games and Economic Behavior*, **53(2)**, 151-169.

Charness, G., L. Rigotti, and A. Rustichini (2007), "Individual Behavior and Group Membership" *American Economic Review*, **97(4)**, 1340-1352.

Coffman, L. (2009), "Intermediation reduces punishment (and reward)," *American Economic Journals – Microeconomics*, forthcoming.

Chaudhuri, A., A. Schotter, and B. Sopher (2009), "Talking ourselves to efficiency: Coordination in inter-generational minimum effort games with private, almost common and common knowledge of advice," *Economic Journal*, **119(534)**, 91-122.

Cialdini, R., R. Reno and C. Kallgren (1990), "A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places," *Journal of Personality and Social Psychology*, **58**(6): 1015-26.

Crawford, S. and E. Ostrom (1995), "A grammar of institutions," *American Political Science Review*, **89(3)**, 582–600.

Deutsch, M. and H. Gerard (1955), "A Study of Normative And Informational Social Influences Upon Individual Judgment," *The Journal of Abnormal and Social Psychology*, **51**(3): 629-36.

Ellingsen, T., M. Johanneson, and J. Mollerstrom (2011), "Social Framing Effects. Preferences or Beliefs?" *manuscript*.

Elster, J. (2009), "Social norms and the explanation of behavior," in P. Hedström and P. Bearman, eds., The Oxford Handbook of Analytical Sociology, Oxford: Oxford University Press, 195-217.

Fehr, E. and U. Fischbacher (2004), "Third-party punishment and social norms," *Evolution and Human Behavior*, **25(1)**, 63–87.

Fehr, E. and S. Gächter (2000a), "Fairness and Retaliation: The Economics of Reciprocity" *The Journal of Economic Perspectives*, **14(3),** 159-181.

Fehr, E. and S. Gächter (2000a), "Cooperation and Punishment in Public Goods Experiments," *American Economic Review*, **90(4)**, 980-994.

Fehr, E. and K. Schmidt (1999), "A theory of fairness, competition, and cooperation," *Quarterly Journal of Economics*, **114(3)**, 817-868.

Frohlich, N. and J. Oppenheimer (1984), "Beyond economic man: Altruism, egalitarianism, and difference maximizing," *Journal of Conflict Resolution*, **28(1)**, 3-24.

Frohlich, N. and J. Oppenheimer (1992), *Choosing justice: An experimental approach to ethical theory*, Berkeley: California University Press.

Gintis, H. (2009), *Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*, Princeton University Press, Princeton: NJ.

Herrmann, B., C. Thoni and S. Gächter (2008), "Antisocial punishment across societies," *Science*, **319**, pp. 1362–1367.

Hoffman, E., K. McCabe and V. Smith (1996), "Social distance and other-regarding behavior in dictator games," *American Economic Review*, **86(3)**, 653-660.

Konow, J. (2000), "Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions," *American Economic Review*, **90(4)**, 1072-1091.

Konow, J. (2003), "Which is the Fairest One of All?: A Positive Analysis of Justice Theories," *Journal of Economic Literature*, **41(4)**, 1186-1237.

Krupka, E. L. and R. A. Weber (2010), "Identifying social norms using coordination games: Why does dictator game sharing vary?", working paper, Carnegie Mellon University.

Lindbeck, A., S. Nyberg and J. Weibull (1999), "Social Norms and Economic Incentives in the Welfare State", *Quarterly Journal of Economics,* **CXIV(1)**, 1-35.

Ostrom, E. (2000), "Collective Action and the Evolution of Social Norms", *The Journal of Economic Perspectives*, **14(3)**, 137-158.

Reuben, E. and A. Riedl (2011), "Enforcement of Contribution Norms in Public Good Games with Heterogeneous Populations", working paper, Maastricht University.

Schotter, A. and B. Sopher (2003), "Social learning and coordination conventions in intergenerational games: An experimental study," *Journal of Political Economy*, **111(3)**, 498-529.

Schotter, A. and B. Sopher (2006), "Trust and trustworthiness in games: An experimental study of intergenerational advice: An experimental study," *Experimental Economics*, **9(2)**, 123-145.

Schotter, A. and B. Sopher (2007), "Advice and behavior in intergenerational ultimatum games: An experimental approach," *Games and Economic Behavior*, **58(2)**, 365-393.

Schwartz, S. 1973. "Normative Explanations for Helping Behavior: A Critique, Proposal and Empirical Test." *Journal of Experimental Social Psychology*, **9**(4): 349-364.

Shogren, J. (1989), "Fairness in bargaining requires a context: An experimental examination of loyalty," *Economics Letters*, **31(4)**, 319-323.

Young, H. P. (2008), "Social norms", in S. N. Durlauf and L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics*, Volume 7, pp. 647–651. Basingstoke: Palgrave Macmillan.

Appendix: Experimental Instructions

This appendix presents a translation of the (Dutch) instructions for the treatment with advice and public payoffs. Places where other treatments differ are given in brackets {…}. The instructions were presented as html-pages. These are separated by horizontal lines, below. Subjects could move from one page to the next (and back) at their own pace.

Welcome

You are about to participate in an experiment. The instructions are simple. If you follow them carefully, you may earn money. Your earnings will be paid to you in euros at the end of the experiment.

Today's payoffs are organized differently than you may be used to at CREED. At the end of the experiment, you will be called forward one by one. We will then publically announce your role today, what you decided, and how much you earned. **Payoffs today are therefore not private and anonymous**.

{In the private payoff case, the previous paragraph was replaced by "At the end of the experiment you will be privately and anonymously paid, one-by-one. Therefore, **no one will know how much you earned** today."}

All monetary amounts are in euros today.

These instructions consist of 5 {4 when there was no advice} pages like this one. While reading them, you may page back and forth by clicking 'next page' or 'previous page' at the bottom of your screen. In some cases, a page may be too large for your monitor. In that case you may use the scroll bar to read through the whole page.

next page

Groups and Roles

At the start of the experiment, we will give each participant **16 euros** as a starting capital.

Then you will be split in groups of three participants. Each group consists of one red player, one green player and one blue player. Your role will be determined completely randomly.

The composition of your group will remain anonymous. You will not know who your co-members are. Others will not know whether or not you are in their group. Similarly, no one will know which role you have. You will not know the role of others.

Only the red players will make a decision today. The green and blue players are **completely dependent** on the red player's choice for their earnings today.

The experiment consists of only one round. Each red player will therefore only be asked to make one decision.

Previous page   next page

Red's decision

The **red** player in a group chooses one of six **possible options.** Each of these options gives specific amounts of money to the red, green and blue player in the group. The options are called **A, B, C, D, E, and F**. The consequences for the players are given in the following table, which will also appear on your screen during the experiment.

	ROOD	GROEN	BLAUW
A	8	0	-16
B	4	4	-12
C	0	8	-8
D	-4	8	-4
E	-8	4	0
F	-12	0	4

[Note: 'rood'=red; 'groen'=green, 'blauw'=blue]

If the **red** player chooses **option A**, **he or she will receive 8 euro**, **green will receive 0 euro and blue will lose 16 euro**.

If the **red** player chooses **option B**, **he or she will receive 4 euro**, **green will receive 4 euro and blue will lose 12 euro**.

If the **red** player chooses **option C**, **he or she will receive 0 euro**, **green will receive 8 euro and blue will lose 8 euro**.

If the **red** player chooses **option D**, **he or she will lose 4 euro**, **green will receive 8 euro and blue will lose 4 euro**.

If the **red** player chooses **option E**, **he or she will lose 8 euro**, **green will receive 4 euro and blue will receive 0 euro**.

If the **red** player chooses **option F**, **he or she will lose 12 euro**, **green will receive 0 euro and blue will receive 4 euro**.

Advice {this page was skipped in the no-advice treatments]

[Note: 'rood'=red; 'groen'=green, 'blauw'=blue]

We imagine that it may be difficult for red to decide what one ought to do in this situation. We will therefore randomly select a few groups and have them **formulate advice**. The groups that give advice will **not subsequently decide from the options**. Each member of an advice group will therefore earn exactly €16 in this experiment.

The advice by every group may be passed on to **more than one group.**
The advice reads:

"**The group that was asked to advise informs you that red ought to choose … in this situation**"

where we substitute the advised option (A, B, C, D, E, F) for the dots.

Each selected groups will separate from the others arrive at an advice in the following way.

1. During **5 minutes (300 seconds)**, the members may exchange thoughts in a **chat box**.
2. Then, each of the three members must **indicate the choice** (A, B, C, D, E, or F) they would like to advise.
3. If there are **three different choices,** the process returns to step 1. There will be a new opportunity to chat. This time, the chat will be for **60 seconds**.
4. If **two or more group members make the same choice**, then that is the advice.

It is not allowed to reveal your identity during the chat.

End

This brings you to the end of these instructions. After everyone has finished, we will start with the experiment.

After you have finished with these instructions, you may indicate so by clicking the 'ready' button at the bottom of your screen. After doing so, you must wait quietly until everyone has finished. This may take a little while, so we ask for your patience.

READY