DISCUSSION PAPER SERIES

IZA DP No. 3860

Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?

Erin Krupka
Roberto Weber

November 2008

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?

Erin Krupka
IZA

Roberto Weber
Carnegie Mellon University

Discussion Paper No. 3860
November 2008

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 3860
November 2008

ABSTRACT

Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?*

We explore the influence of social norms on behavior. To do so, we introduce a method for identifying norms, based on the property that social norms reflect social consensus regarding the appropriateness of different possible behaviors. We demonstrate that the norms we elicit, along with a simple model combining concern for norm-compliance with utility for money, predict changes in behavior across several variants of the dictator game in which behavior changes substantially following the introduction of minor contextual variations. Our findings indicate that people care not just about monetary payoffs but also care about the social appropriateness of any action they take. Our work also suggests that a social norm is not always a single action that should or should not be taken, but rather a profile of varying degrees of social appropriateness for different available actions.

JEL Classification: C91, C72, D34

Keywords: norms, matching games, dictator games

Corresponding author:

Erin Krupka
IZA
P.O. Box 7240
D-53072 Bonn
Germany
E-mail: ekrupka@iza.org

* The authors thank the Ford Foundation and Carnegie Mellon University's Center for Behavioral Decision Research for financial support. We thank Nicholas Bardsley, Cristina Bicchieri, Ted O'Donoghue, Tanga McDaniel, and participants at several workshops, seminars, and conferences for helpful comments and suggestions.

I. Introduction

Social norms have long been recognized as an important influence on behavior in other social sciences such as social psychology (Sherif 1936; Cialdini et al. 1990) and sociology (Merton 1957; Coleman 1990). However, in economics social norms have received significant attention only relatively recently, mainly as a tool for explaining seemingly anomalous behavior such as involuntary unemployment (Akerlof 1980), conformity (Bernheim 1994), costly punishment (Fehr and Gächter 2000), tipping (Conlin et al. 2003), and macroeconomic phenomena such as why consumption may track income even when wealth levels remain unaffected (Akerlof 2006).

One possible reason for the relative absence of social norms in economic research is that they are difficult to measure or quantify and it is therefore difficult to predict the precise influence they will exert on behavior. As a result, social norms are usually incorporated into economic research as *post hoc* interpretations for behavior or outcomes that are otherwise difficult to explain (Cole et al. 1992; Nee 1998; Bowles 1998; Fehr and Gächter 2000; Ostrom 2000), and they are identified primarily by measuring behaviors that are theoretically related to the norm (Lindbeck et al. 1999; Nyborg and Rege 2003; Fehr and Fischbacher 2004; Fehr and Camerer 2004). Because norms are usually studied indirectly in economics,[1] they are rarely used to form precise predictions about behavior.[2]

[1] A few experimental studies manipulate the likely presence or strength of a social norm in the laboratory by varying features of the choice context (Berg et al 1995; Ross and Ward 1996; Krupka and Weber 2007; Andreoni and Bernheim 2008) and demonstrate resulting changes in behavior that are consistent with the influence of a norm. Our experiment employs a similar procedure, by varying a feature of the environment to study a social norm and its impact on behavior. However, instead of varying the contextual features and then examining their indirect effect on behavior (through the norm, or perhaps through increased concern for complying), we directly elicit the norm and show that it is changing social norms that are responsible for the observed changes in behavior.

[2] Andreoni and Bernheim (2008) nicely demonstrate how varying features of a choice environment, such as perceived responsibility for outcomes, produce changes in behavior consistent with the predictions of a model of norms in which individuals care about how others perceive their degree of norm-compliance. Unlike our paper, they begin by taking the social norm as given, at a 50-50 division of wealth. While we also find that 50-50 splits are

In this paper, we aim to put the horse (norm) before the cart (behavior), by measuring a social norm and using it to predict behavior *a priori*. We do so in the context of other-regarding behavior in the "dictator game," where recent laboratory experiments demonstrate that minor contextual features of a choice environment lead to substantially different choices and outcomes.[3] We show that such changes in behavior are entirely consistent with varying social norms.

We begin by defining social norms, and presenting a generic model of utility over outcomes and norm compliance. We then demonstrate how one can use simple coordination games to identify social norms. Using the model, we show how the social norms we elicit with the coordination games yield precise and testable behavioral predictions, which we compare to data.

We first elicit social norms governing behavior in two theoretically equivalent, but superficially dissimilar, contexts involving other-regarding behavior (i.e., variants of the dictator game) and show that the elicited norms accurately predict changes in behavior of laboratory subjects between these two contexts. We also elicit social norms governing behavior in related previous experiments in economics, and find that the observed sensitivity of behavior to the experimental treatments is entirely consistent with the social norms that we elicit. In the conclusion, we discuss related work that validates the norm elicitation method, by showing how

---

generally viewed as most "norm-compliant," thereby confirming the broad appeal of 50-50 norms, we also demonstrate that part of the influence of social norms on behavior comes from social perceptions of the appropriateness of other actions, relative to the 50-50 split. Thus, while Experiment 2 in this paper keeps the key features of their model constant, we find significant changes in behavior that are consistent with our interpretation of the influence of norms.

[3] This apparent "instability" in behavior has led some researchers to question the value of generalizing from such laboratory experiments to the field (Levitt and List 2007; List 2007). Our work at least partially addresses this concern, by demonstrating how such behavior corresponds to varying and identifiable social norms. The sensitivity of behavior in the laboratory to the context of the experiment can be interpreted in a manner similar to how behavior in the field is sensitive to context. For these laboratory experiments we demonstrate that such sensitivity may be explained once varying social norms are identified. Therefore, the reason why someone might share in one dictator experiment but not in another very similar one might be the same as why one tips at a coffee shop but not at a fast-food restaurant, or in the U.S. but not in Europe.

elicited norms over non-laboratory behavior correspond to well-known and externally verified social norms.

II. Defining and Identifying Social Norms

Following Elster (1989), we note two important features of social norms. First, social norms are typically not outcome-oriented, but instead directly govern behaviors or actions. As Elster notes, "The simplest social norms are of the type: Do X, or: Don't do X." (p. 99). Second, the "social" element of norms requires that they be jointly recognized, or mutually held, by members of a population.[4] These two features – that social norms typically apply to actions rather than outcomes and that they must be jointly recognized – are present in most researchers' definitions (Bettenhausen and Murnighan 1991; Fehr and Gächter 2000; Bicchieri 2006). For example, Ostrom (2000) defines social norms as "*shared understandings* about *actions* that are obligatory, permitted, or forbidden" (pp. 143-144, emphasis added).

Further, we distinguish norms regarding what one "ought" to do, or injunctive norms, from customs or actions that people regularly take, or descriptive norms (Deutsch and Gerard 1995; Bicchieri 2006). Both kinds of norms influence behavior (Cialdini et al. 1990; Krupka & Weber 2007; Bicchieri & Xiao forthcoming). However, our focus here is on injunctive social norms, i.e., those described by Elster as prescribing what one "should do" or "should not do." As we will show, social norms concerning the appropriateness of behavior (injunctive norms) are sufficient for explaining a considerable amount of other-regarding behavior.

[4] At least implicitly, most definitions distinguish between social norms and personal norms. The former, which are our focus here, usually refer to a common understanding among members of a group. An individual member of a group has a belief that others in the group judge a particular behavior appropriate (or inappropriate) and that the others in the group assume the individual is aware of this judgment. In this sense, the individual and the group *share an understanding regarding the in/appropriateness of behavior* and this shared understanding is a social norm (cf. Bicchieri 2006; Young 2008).

Therefore, we define social norms as shared perceptions, among members of a population, regarding the appropriateness of different behaviors. They are things that people in the population jointly recognize one should or should not do, and people who belong to the population expect others to be aware of and understand this agreement. The power of social norms comes both from the willingness of people within the population to punish (or reward) others' adherence to (or deviation from) them and from the experience of positive or negative emotions produced by one's own adherence or deviation from a social norm (Elster 1989; Fehr and Gächter 2000).

To formalize our definition, let $A = \{a_1, \ldots, a_K\}$ represent a set of K actions available to an individual. A social norm, $N(a_k)$, assigns to each action a degree of appropriateness ($N(a_k) > 0$) or inappropriateness ($N(a_k) < 0$). Therefore, we assume that if for an action, a_k, there is joint recognition that the action constitutes "appropriate," or socially prescribed, behavior, $N(a_k) > 0$, while if there is joint recognition that an action constitutes "inappropriate," or socially proscribed, behavior, $N(a_k) < 0$.[5]

An important feature of the above definition is that a norm does not necessarily constitute a binary classification, such that a particular action (the "norm") should be taken, by assumption leaving all remaining actions as those that should not be taken.[6] Instead, our definition of a social norm applies to the entire set of possible actions, and allows actions to vary in the degree to which they satisfy social norms. Therefore, for example, while there may be social agreement

[5] This definition assumes a single, homogenous population that comprises both the individual making the decision and the social group determining perceptions of appropriateness. In some cases, where individuals interact within relatively homogenous populations, this seems a reasonable way to study social norms. However, there may also be situations in which it is important to define multiple reference groups for which different particular social norms hold. For example, social norms of punctuality and tipping may differ significantly between two populations (see Krupka et al. 2008). In such instances, we can define $N(a_k)$ in reference to a particular group (for example, by letting $N_g(a_k)$ denote the social norms for group g) and recognize that such social norms are only likely to impact an individual when she perceives herself to be a member of the group.

[6] Such a definition would be possible in our framework by, for example, assigning $N(a_k) > 0$ to only one action (the "norm") and letting all other actions have a constant value of $N(a_k) < 0$.

that it is always appropriate to arrive on time in many Western cultures, there may be some instances in which arriving 5 minutes late is less socially inappropriate (meeting friends at a bar) than others (arriving at a funeral). We will also see this to be important in our analysis of experimental data, in that the most socially appropriate behavior is the same across contexts (share the experimental endowment equally (cf. Andreoni and Bernheim, 2008)), but where deviations in the relative appropriateness of the other actions are important influences on behavior.

The individual cares about both the payoff produced by the selected action, $\pi(a_k)$, and the degree to which the action is socially appropriate:

$$u(a_k) = V(\pi(a_k)) + \gamma N(a_k). \quad (1)$$

The function *V()* represents the value the individual places on the monetary payoff, and we assume only that this function is monotonically increasing. The parameter $\gamma \geq 0$ represents the degree to which the individual cares about adhering to social norms.[7] An individual entirely unconcerned with social norms ($\gamma = 0$) will always select the payoff-maximizing action. On the other hand, as γ_i increases, an individual will derive greater utility from selecting actions that are socially appropriate and less utility from those that are not.[8]

It follows directly from the above model that behavior may change substantially across choice environments in which the sets of payoffs are identical. For example, consider two choice environments, $A = \{a_1, a_2\}$ and $A' = \{a_1', a_2'\}$, such that $\pi(a_1) = \pi(a_1') > \pi(a_2) = \pi(a_2')$. Then, if there exist no social norms in either environment ($N(a_k) = N(a_k') = 0$, for k = 1, 2) or if an

[7] Other researchers have noted that individuals care heterogeneously about norm compliance (Ostrom 2000, Fisher and Huddart 2008, Andreoni and Bernheim 2008). Such heterogeneity is also common in most models of social preferences (e.g., Fehr and Schmidt 1999; Andreoni and Miller 2002).

[8] Cases in which $\gamma_i < 0$, which we do not explore here, might correspond to individuals who are anti-social, or derive utility from violating norms.

individual does not care about social norms ($\gamma = 0$), the individual will choose a_1 in the first environment and a_1' in the second. However, if social norms differ between the two choice environments, the individual may select actions corresponding to different payoffs in the two environments. For example, if $N(a_1) = N(a_1')$ but $N(a_2) < N(a_2')$, then for large enough values of γ an individual will select a_1 in the first environment and a_2' in the second environment.

While the above formal definition and analysis of social norms is straightforward and simple, it presents a useful framework for understanding how behavior might change even when choice environments are payoff-equivalent. It also provides a testable relationship between the degree of social approval of actions ($N(a_k)$) and individuals' willingness to take those actions, provided one has a reasonable method for capturing the "social appropriateness" of the different available actions.

In the rest of this paper, we attempt to predict and explain behavior using elicited measures of social appropriateness. We study choices in several environments in which individuals unilaterally choose allocations of wealth between themselves and others, as in the well-known dictator game (Hoffman et al. 1994; see Camerer 2003).

We measure the extent to which actions are socially appropriate or inappropriate using a novel elicitation method. We elicit social norms over possible action choices across different contexts, from individuals that are distinct from those making choices in those contexts. To do so, we present respondents with a description of a choice environment, including all the possible available actions. We ask respondents to judge the social appropriateness of *each* action on a four point scale that ranges over "very socially inappropriate", "somewhat socially

inappropriate", "somewhat socially appropriate" to "very socially appropriate."[9] Finally, we provide respondents with incentives not to reveal their own personal preferences but instead to match the responses of others. Thus, respondents play a "pure matching" coordination game (Schelling 1960; Mehta, Starmer & Sugden 1994) in which their goal is to anticipate the extent to which others will rate an action as socially appropriate or inappropriate, and to respond accordingly.

Because social norms reflect "shared understanding," coordination games present a reasonable way to identify such socially-held judgments. From a game-theoretic point of view, matching games such as the one we use in our experiment, have a number of equilibria and nothing intrinsic to the game makes one equilibrium favored (or focal) over the other. Schelling (1960) theorized and Mehta et al. (1994) and Sugden (1995) demonstrated that prominence derived from common culture and shared experiences can create focal points. In our experiment, we allow jointly recognized social norms to create focal points in the matching game. That is, if there is general social agreement that some actions are more or less socially appropriate, respondents attempting to match others' responses are likely to rely on such shared perceptions to help them do so. Therefore, we predict that responses will capture not personal perceptions of the appropriateness of behaviors, but instead will capture jointly recognized perceptions of appropriateness, which we have defined to be social norms.[10]

Our experiments involve first identifying social norms and using them to predict and explain behavior. We begin by focusing on two payoff-identical variants of the dictator game.

[9] The decision to have only four appropriateness categories was made after considering the tradeoff between having too few (in which case it would be harder to discriminate between degrees of appropriateness) and having too many (in which case it might be too difficult for subjects to match on the social norm, perhaps leading them to attempt to match using other focal principles). Further, we omitted the "neutral" category as this would have been a focal point separate from the focal point stemming from the social norm.

[10] Many previous researchers have noted the important relationship between social norms and equilibrium selection in games (Kandori 1991; Young 1998).

We first describe these two choice environments and elicit social norms over behavior in these two environments from a group of subjects who never actually play these games (Experiment 1). We then use the elicited social norms from Experiment 1 to predict how behavior changes across the two environments and test these predicted effects of norms using data collected from a second, separate, group of subjects who made actual choices in one of the two environments (Experiment 2).

A second part of our analysis involves identifying social norms governing behavior in four previously studied other variants of the dictator game (Dana et al. 2007; Lazear et al. 2006; List 2007; and Bardsley 2008). Therefore, as part of Experiment 1, we also use our elicitation method to measure the degree of social appropriateness of different actions available in these particular experiments. We demonstrate that the identified social norms explain considerable variation across treatments in both our own experiment (Experiment 2) and across these previously studied experiments.

III. Identifying social norms in payoff-equivalent environments (Experiment 1)

Consider the following two choice environments. In a "standard" dictator game, an individual initially receives $10 while another person receives $0. The individual must then decide how much, between $0 and $10 in one dollar increments, to share with the other person. In a "bully" variant of the game, the individual and other person both initially receive $5 and the individual can give or take any amount between $0 and $5, again in one-dollar increments, to or from the other person. Both choice environments offer the deciding individual 11 choices over wealth allocations ranging from ($10, $0) to ($0, $10), but vary in the actions required to obtain those dollar allocations.

While the two choice sets are identical in terms of final payoffs, they differ in contextual features of the actions required to achieve those payoffs. In the standard case, any outcome other than ($10, $0) involves "giving" money to the other person, while in the bully variant all outcomes from ($10, $0) to ($6, $4) involve the individual "taking" from the other person. Therefore, it is possible that social norms governing the two sets of behaviors might differ considerably, even though the resulting outcomes do not. In particular, we expect social norms to differ over actions that involve "taking" vs. "giving," holding the resulting payoffs constant.

To identify social norms in the two choice environments, we applied our elicitation method to obtain ratings of the extent to which different actions in the two environments are socially appropriate or inappropriate. Respondents received incentives to match the modal response provided by others.

III.A. Experimental Design for Experiment 1

115 subjects were recruited from a list containing Carnegie Mellon and University of Pittsburgh students and employees. The experiments were conducted at the Pittsburgh Experimental Economics Laboratory (PEEL) at the University of Pittsburgh. Participants received $7 for showing up to the experiment, as well as any money they accumulated during the experiment, in cash. Subject payment in the matching task was not tied to the hypothetical dictator games about which they read.

Participants received instructions and randomly-assigned cards containing ID numbers from 1 to 20. The instructions (see Appendix) explained that they would read descriptions of different situations in which a person ("Individual A") faced a choice among several possible

alternatives. For each situation, they would rate the extent to which each alternative available to the person was "'socially appropriate' and 'consistent with moral or proper social behavior' or 'socially inappropriate' and 'inconsistent with moral or proper social behavior.'"

Participants then read, as an example, about a hypothetical situation.[11] The instructions then showed participants how they might indicate responses for such a situation. Participants were then told that one of the situations for which they were to provide appropriateness ratings would be selected at random at the end of the session, and that one of the possible action choices in this situation would also be randomly selected. If, for this action choice, the participant's appropriateness rating was the same as the modal response in the session, then that participant would receive an additional $5 payment at the conclusion of the session.

Subjects then saw a description of either the standard or bully variants of the dictator game. Subjects who rated the choice actions never actually played this game, but only read about the situation and were asked to consider all of the actions that A (the dictator) could take. In each session, only one of these two variants, standard or bully, was used. Thus, no subject read descriptions of both the bully and standard choice contexts. The description of the situation stated that the target individual (Individual A) was matched with another random and anonymous person (Individual B) and that both people in the hypothetical situation would receive a "small participation fee" as well as any money produced by Individual A's choice.

The description then listed the eleven action choices available to Individual A. The labels associated with these action choices varied depending on which dictator game variant subjects were asked to consider (either standard or bully). Subjects were also shown the monetary payments to each individual (A and B) produced by every listed action choice. For each possible

[11] In the example, the target individual found a wallet at a coffee shop and faced four alternatives: taking the wallet, asking others if the wallet belonged to them, leaving the wallet alone, or giving it to the store manager.

action choice available to Individual A, a subject had to rate the choice as either "very socially inappropriate," "somewhat socially inappropriate," "somewhat socially appropriate," or "very socially inappropriate," with the goal of matching this rating to the modal response in the session. That is, for each possible choice available to Individual A, the subject had to record his or her best guess of the modal appropriateness judgment in the session.

After completing their ratings of social appropriateness for all actions available in either the standard or bully variant of the dictator game, subjects then saw descriptions of four additional variants of the dictator game. Each variant corresponded to a variant of the dictator game used in previous experimental research (Dana et al., 2007; Lazear et al 2006; List 2007; Bardsley 2008). We discuss these variants in more detail in Section V.

After subjects had indicated social appropriateness ratings for all possible choices in all five choice scenarios, the experimenter randomly selected one scenario and one possible choice in that scenario. The experimenter computed the modal response for that choice and informed subjects of whether or not their appropriateness rating had matched the modal rating. Subjects were then paid privately, receiving a $7 participation fee and an additional $5 if they had selected the modal appropriateness rating for the selected scenario.

III.V. Results of Experiment 1

Recall that we anticipated that "taking" would generally be considered less socially appropriate than "giving," even when they produced identical outcomes. Therefore, we expected that for those wealth allocations that left the dictator (Individual A) with more money than the recipient, the corresponding actions would be generally considered less socially appropriate in the bully variant than in the standard variant of the dictator game.

We converted subjects' responses into numerical scores. A rating of "very socially inappropriate" received a score of -1, "somewhat socially in appropriate" a score of -1/3, "somewhat socially appropriate" a score of 1/3, and "very socially appropriate" a score of 1. Table 1 presents summaries of subjects' social appropriateness ratings by condition. Each row corresponds to one possible action choice that Individual A could take and is also denoted by the final wealth distribution produced by that action choice in the first column (payoff for A, payoff for B). For each of the two variants, the next several columns report first the mean of the social appropriateness ratings (ranging from complete agreement on "very socially inappropriate" (-1) to complete agreement on "very socially appropriate" (+1)), and then the full distribution of responses. The final column reports the result of a Wilcoxon rank-sum test, a non-parametric comparison of the two distributions of responses that accounts for their ordinal nature.

Not surprisingly, the general pattern of social appropriateness ratings is the same across the two choice environments. There is substantial social agreement that maximizing A's own payoff and leaving the other person with nothing ($10, $0) is very inappropriate in either variant. Similarly, subjects recognize large social agreement that the action that produces equal payoffs ($5, $5) is very socially appropriate in either environment.

In addition, actions that leave the recipient with more money (($4, $6) to ($0, $10)) are much more ambiguous, though the modal and median responses generally lie between "very" and "somewhat" socially inappropriate. Still, a significant proportion of respondents rate such behavior as socially inappropriate, and this proportion generally increases with other-regarding inequality.[12]

[12] This might reflect the belief that it is socially inappropriate to be "too generous" – for example, when one gives a gift that is too expensive or when one attempts to tip a member of a profession that generally does not accept tips.

While in the above instances the social appropriateness ratings generally agree across the two environments, we see a different pattern for outcomes in which the dictator obtains most, but not all, of the wealth. Recall that for A to obtain between $6 and $10 in the standard environment, the dictator must "give" to the other person while the bully environment requires the dictator to "take" from the other person. The ratings confirm our hypothesis that "giving" is more socially appropriate than "taking." For every outcome from ($9, $1) to ($6, $4), the mean rating for the corresponding action is higher in the standard (giving) environment than in the bully (taking) environment. And for every one of these outcomes the differences in ratings is statistically significant.

Subjects are still quite able to anticipate others' ratings – the modal response almost always receives over half of the responses – but what they agree upon often differs substantially. For example, for the wealth allocation ($8, $2), the modal response in the standard environment for giving $2 to the other person is "somewhat inappropriate" (52%). But in the bully environment, where the same outcome involves taking $3 from the other person, there is social agreement that the action is "very inappropriate" (57%). Similarly, for the wealth allocation ($6, $4), the modal response in the standard environment is "somewhat appropriate" (66%), but in the bully environment it is "somewhat inappropriate" (49%).

III.C. Behavioral Predictions

The model in equation 1 and the elicited norm ratings $N(a_k)$ in Table 1 lead directly to two main predictions regarding how behavior will differ between the two environments. To generate the predictions, we compare the behavior of two identical populations of agents, each confronted with a different version of the decision (bully vs. standard). We assume that all agents

place the same value on monetary earnings, meaning that *V()* is identical for all agents. Each population contains a continuum of agents, characterized by the degree to which they care about norms – meaning that we allow heterogeneous concern about norm compliance (γ_i).[13] To be strict, the predictions require that the range of γ_i be sufficiently wide to rule out situations in which all agents in both populations make identical decisions.[14]

Prediction 1: *More agents will select the action producing the equal-split ($5, $5) allocation in the bully environment than in the standard environment.*

Prediction 2: *Conditional on not selecting the action producing the equal-split ($5, $5) allocation, more agents will select the action producing the payoff-maximizing ($10, $0) allocation in the bully environment than in the standard environment.*

The two predictions are straightforward. The first prediction follows from the fact that the loss in social appropriateness from moving from the equal-split action to any action yielding a greater payoff for the decision maker is relatively greater in the bully treatment than in the standard environment. As an example, from Table 1 we see that the difference between appropriateness ratings for keeping $9 and keeping $5 in the standard dictator game is 1.61 (= 0.90 – (-0.71)), meaning that an individual in the standard game will prefer sharing $5 to sharing

[13] The two predictions follow directly from the model and the differences in ratings. Note first that agents should never allocate themselves less than half of the wealth because it produces both a lower payoff and lower social appropriateness than choosing the action that produces ($5, $5). For the first prediction, note that in the bully variant every allocation that produces more wealth than ($5, $5) for the agent yields a lower social appropriateness rating than in the standard dictator game. Therefore, selecting the action corresponding to the equal split will be more attractive relative to every other feasible choice in the bully variant than in the standard game. A similar argument applies when one considers only actions that produce favorable inequality – in these cases leaving the recipient with nothing becomes relatively more attractive than all other options. Further details are available upon request from the authors.

[14] For example, if all agents place a sufficiently high weight on norm-adherence (i.e., if all γ_i are very high), then the equal-split outcome would be predicted in both choice environments.

\$1 when $\gamma_i > {V(9)-V(5)}/{1.61}$. However, in the bully game, the corresponding difference is larger (1.80), meaning that preferring to take nothing over taking \$4 in the bully game occurs when $\gamma_i > {V(9)-V(5)}/{1.80}$. Thus, some individuals may prefer implementing the (\$9,\$1) allocation over the (\$5,\$5) allocation in the standard dictator game, but may have the opposite preferences in the bully variant. This is true for all comparisons of utility between the equal split allocation and an allocation that yields more wealth for the dictator. Some individuals find the equal-split allocation more attractive in the bully variant than in the standard game, relative to other feasible choices, and the opposite is never true.

Similarly, conditional on not selecting the equal-split action, implementing the payoff-maximizing (\$10,\$0) action is relatively more socially appropriate – compared to actions that produce payoffs between (\$9,\$1) and (\$6,\$4) – in the bully environment than in the standard one. Thus, in the bully environment, individuals are less likely to select something other than the equal split (Prediction 1), but if they do then they are more likely to take all the wealth (Prediction 2).

The above shows how we can combine a simple model of utility, which includes both utility from monetary payoffs and from complying with social norms, with the appropriateness measures we obtain from the coordination game. This approach yields novel predictions regarding how behavior should change between the two environments, which we now test.

IV. Testing the model's predictions against behavioral data (Experiment 2)

To evaluate the accuracy of the above behavioral predictions, we conducted an experiment that placed a different set of subjects in one of the two choice environments described in Experiment 1. These subjects, who had not participated in the coordination games

used to elicit social norms in Experiment 1, played either the standard or bully version of the dictator game for actual monetary payoffs. The difference between the two environments was whether one subject in a pair received $10 and chose how much to give to the other subject (standard) or whether both subjects received $5 and one subject chose how much to give to or take from the other (bully). The possible set of final payoff allocations was identical in both environments.[15]

IV. A. Experimental Design for Experiment 2

Our experiment took place at the end of several large lecture classes at Carnegie Mellon University. We recruited participants by asking for up to 30 volunteers to remain after class for a 5 minute decision making experiment.[16] Sessions consisted of between 16 and 30 participants. Participants received $2 for participating, in addition to any money from the allocation choices made in the dictator game variants described below.

Once all non-participants left the classroom we divided participants into two groups seated in different areas. One group (dictators) received instructions, which were also read aloud so that the other group (recipients) could hear the instructions.

In the *standard* dictator game, each dictator received a yellow envelope labeled "money for you" that contained ten $1 bills. The other group (recipients) received empty white envelopes labeled "money for other person." Instructions were read aloud describing the dictator choice, in which dictators would make a (double-blind) anonymous and private decision of how much of the $10 in their envelope to share with the paired recipient.

[15] Our experimental design joins a group of experiments that examine the effect of varying initial endowment levels such that dictators may "take" money from player 2 (Eichenberger and Oberholzer-Gee 1998, Swope et al. 2008, Cox et al. 2007, Bardsley 2008)

[16] We recruited from six science and math classes. While we did not collect gender information, our sample was predominantly male.

After instructions were read aloud, one experimenter collected the empty white envelopes from recipients and waited by the door to the hallway outside the room. Dictators exited the room one at a time, and each dictator received one recipients' empty white envelop prior to exiting (the dictator already had his or her yellow envelope in hand). Outside the room, in the hallway, dictators found a large sealed box with an open slit at the top. As described in the instructions, the dictator privately allocated money between the two envelopes, then placed the white envelope labeled "money for other person," which had originally been empty, inside the box, and left with whatever remained in the yellow envelope. This procedure allowed decisions to be anonymous.[17] This concluded the experiment for the dictator.

Once all dictators had left, one of the experimenters then brought the box back into the classroom, where the recipients had been instructed to form a line. As each recipient stepped up to the experimenter table, one at a time, an experimenter opened each white envelope, counted the number of $1 bills aloud, and handed the bills to the recipient. The other experimenter recorded the amount received by the recipient. This concluded the experiment.[18]

In the *bully* variant of the dictator game, procedures were identical except that the two envelopes handed out at the beginning of the experiment each contained five $1 bills. The instructions informed dictators that they would be able to give up to $5 to or take up to $5 from the other person (instructions are included in the Appendix).

IV. B. Results of Experiment 2

[17] The box was placed in such a way that the experimenter standing at the door could see part of the back of the person standing at the box (but not enough to be able to determine whether the subject was reallocating money between the envelopes or when the envelope was being placed in the box). The experimenter could observe the subject departing from the box area, which allowed the experimenter to know when to send the next dictator out of the classroom. This minimal observation also prevented subjects from being able to open the box.

[18] For accounting purposes, and to maintain anonymity of actions, dictators signed a sheet stating that they received a $2 participation fee as well as $10 to allocate between themselves and another participant. Recipients conversely signed a sheet stating that they received $2 and may have also received some money from another participant.

Figure 1 presents the results across the two conditions. There were 37 dictators (74 subjects) in the standard treatment and 54 dictators (108 subjects) in the bully treatment.

The mean amount allocated to the recipient was $2.62 in the standard game and $2.88 in the bully treatment. The results in the standard treatment are quite similar to those in other dictator games: subjects share about 25% of the endowment and most dictators make some non-zero transfer (Camerer 2003, Forsythe et al. 1994).[19]

Table 2 presents statistical tests of the changes in behavior across the two treatments. The first two models demonstrate that more is shared with the recipient in the bully treatment, relative to the standard but that the difference is only marginally significant (p=0.09). We include as a control variable the size of the class from which students were recruited – class size ranged from 87 to 184 – since this is potentially a measure of social distance (Bohnet and Frey 1999). As we expected, class size is significantly negatively related to amount shared.

We now test the behavioral predictions generated by the model when it is combined with the norm elicitation results in Experiment 1. The first prediction was that more participants would select the action corresponding to the equal-split allocation ($5, $5) in the bully treatment than in the standard treatment. Figure 1 reveals strong support for this prediction. If we exclude those subjects who shared more than $5,[20] then in the standard condition 6 of the remaining 33 participants (18 percent) gave $5 to the recipient. In the bully treatment, however, the proportion is much higher: of 49 participants (excluding those who shared more than $5), 18 (37 percent)

[19] The amount shared is higher than in other experiments with this high level of anonymity (Hoffman et al. 1994 and 1996). However, the social distance between dictators and recipients in our experiment is probably lower than in typical studies, as they are classmates, and social distance is negatively related to sharing (Bohnet and Frey 1999).
[20] Such behavior is usually present, though rare, in most dictator experiments (Camerer 2003), and is inconsistent with both our model and most other models of social preferences (Fehr and Schmidt 1999, Bolton and Ockenfels 2000). Our results do not substantively change if we include those 9 participants in the analysis.

neither took from nor gave money to the recipient, thus producing a final allocation of ($5, $5). As column 2 in Table 2 reveals, this difference in behavior is statistically significant ($p < 0.01$).

The second prediction deals with what subjects do if they do not select the equal-split action. We predicted that, conditional on allocating less than $5 to the recipient, more dictators would leave the recipient with $0 in the bully variant than in the standard game. In the standard game, 27 participants gave less than $5 to the recipient, and of these 11 (41 percent) gave $0. In the bully variant, 31 participants took money from the recipient, and of these 16 (52 percent) left the recipient with $0. The percentage is higher in the bully variant than in the standard treatment, as we predicted, though this difference is not large. As Table 2 reveals, this difference is marginally statistically significant using a two-tailed test ($p = 0.10$).

Experiment 2 demonstrates that behavior changes significantly across two payoff-equivalent choice environments. In both the bully and standard treatments, subjects choose how much of $10 to keep for themselves, and how much to allocate to an anonymous recipient. However, subjects allocate significantly more to the other person when some of the actions involve “taking from,” rather than “giving to,” the other party.

We also find that the changes in behavior are generally consistent with the predictions of our simple model and the elicited ratings of social appropriateness from Experiment 1. Significantly fewer dictators are willing to choose actions that involve “taking” in the bully variant, than those who are willing to choose comparable “giving” actions in the standard treatment. This is because the taking actions are generally regarded as more socially inappropriate than comparable “giving” actions, even when the two sets of actions produce identical outcomes.

We interpret changes in behavior across the two environments through changes in the social appropriateness of seemingly identical – in terms of payoffs – actions. However, to test our interpretation of why sharing varies in dictator games, it is important to demonstrate that changes across a larger set of choice contexts are consistent with our model. We now turn to an analysis of previously-collected data on variants of the dictator game.

V. Re-analyzing previously-collected dictator game data

In Experiment 1, after providing ratings of the social appropriateness of actions in either the standard or bully environment, all 115 subjects also performed similar ratings for four other variants of the dictator game. These variants were all modeled closely after actual experiments conducted in previous research. The task of providing ratings was presented in exactly the same format as for the standard and bully variants – subjects saw a list of the possible actions available to an Individual A (the dictator) in that particular experimental treatment, and then attempted to coordinate ratings of social appropriateness for each possible action with other subjects in the session.[21] Participants' incentives were identical to those for the first variant (standard or bully) that they had encountered.

We now show that the elicited ratings are consistent with behavior in these experiments. In each case, we first describe the "puzzling" results produced by the particular variant of the dictator game, and then show that the elicited social appropriateness ratings are consistent with such behavior.

[21] No feedback was provided until subjects had completed the entire experiment. That is, subjects first played a matching game on either the standard or bully variant of the dictator game, then completed matching games on four other games (the order of which was randomized). After subjects had completed the experiment, and all responses were collected, subjects received feedback about others' choices in one scenario.

V. A. Dictator game with a sorting option:

Lazear et al. (2006) explored a variant of the dictator game in which subjects could opt to not play the game (by "passing"), in which case the dictator received $10 and the other participant received $0 without learning that a dictator game could have been played.[22] The introduction of this option, which replicates the payoffs produced from sharing nothing, has a strong effect on sharing. In Lazear et al.'s Experiment 1, mean sharing decreases from €1.87 in the treatment without a sorting option to €0.58 in the treatment with a sorting option. This is largely the result of a majority of dictators in the sorting treatment selecting not to play the game (72 percent).

To see why the sorting option is so frequently chosen (even among people who share positive amounts when no such option is available) and why behavior changes so significantly between the environments with and without sorting, we consider the ratings of social appropriateness given to actions in the two environments. Figure 2 presents the average ratings, for each action (represented in terms of final payoffs, on the x-axis), of social appropriateness both from the standard version of the dictator game and the variant with the additional ($10,$0) sorting option.

The dark line presents the ratings (from Table 1) for the standard version of the dictator game. The lighter line presents the ratings from the sorting variant, in cases where the dictator chose to play the game. The two lines are very close, indicating that social appropriateness ratings differed very little for actions in the dictator game depending on whether subjects were required to play the game (standard) or had the option of not playing the game but then chose to

[22] This kind of game has also been studied by Dana et al. (2006), who include a ($9, $0) sorting option (Lazear et al. also include some sorting treatments in which the aggregate endowment does not sum to $10). We focus on Lazear et al.'s Experiment 1, because we are primarily interested in changes in behavior when the outcomes among which an individual is choosing produce identical payoffs, and ($10, $0) is an available outcome in the dictator game (but ($9,$0) is not).

play (sorting, play). In all comparisons between actions producing identical outcomes, conditional on the dictator playing the game, the differences in ratings are statistically insignificant ($z < 1.11$ using a rank-sum test for every comparison).

The grey square corresponds to the mean rating given to the choice of taking the sorting option and *not* playing the game ("Sorting(Pass)"). This action implements a ($10,$0) outcome, with the recipient remaining uninformed about the game. Thus, the payoffs produced are identical to those from playing the game and keeping all the money. However, as the ratings reveal, choosing the sorting option is considered far less socially inappropriate. The mean rating is 0.08 for Sorting(Pass) versus -0.86 for keeping all $10 in the standard game and -0.83 for choosing to play and keeping $10 in the sorting variant ("Sorting(Play,$10)"). The differences in appropriateness ratings between the Sorting(Pass) option and keeping $10 when playing the game are highly statistically significant ($p < 0.001$) both for the standard game ($z = 8.61$) and for the sorting variant ($z = 10.02$).[23]

To see how considerations of social appropriateness are likely to influence outcomes between the two variants, consider the relatively high rating (0.08) given to the action choice "pass." According to our model, this option is desirable relative to other actions, as it produces the highest possible monetary payoff (i.e., $V(\pi(a_k))$ is maximized) at a relatively low cost in terms of disutility form social disapproval ($N(a_k)$). Thus, people who select to share positive amounts in the standard variant of the dictator game might prefer the sorting option, which provides a large monetary payoff with less disutility from violating social norms than would obtaining the same payoff when playing the dictator game.

[23] Interestingly, there is less agreement regarding the social appropriateness of choosing not to play the dictator game (and keeping $10) than there is for other choices in either variant. The modal rating (32 percent) is "somewhat inappropriate," but there are high frequencies of other responses as well (27 percent for "somewhat appropriate").

For example, consider an individual for whom the marginal utility from a monetary payoff of \$10 is *k* times the marginal utility from a payoff of \$5 (i.e., $V(\$10) = kV(\$5)$). Then, in the standard variant this individual will prefer the equal-split allocation over keeping all the money whenever $\gamma_i > (k-1)V(\$5)/1.76$ (this follows from Table 1).[24] However, in order to prefer the "pass" option in the sorting variant of the game over playing the game and selecting the equal split ("Sorting(Play,\$5)"), it is only necessary that $(k-1)V(\$5)/0.82 > \gamma_i$. Therefore, for a range of concern for social norms (γ_i) subjects might be willing to prefer sharing \$5 in the standard dictator game over keeping everything, but prefer to opt out of playing the game altogether.

V B. Dictator game with additional taking options:

We also explored another variant of the dictator game in which the payoff space is extended to include additional options in which the dictator can take money from the recipient. In the "Take-2" variant, we described a dictator game in which dictators were allowed to give between \$0 and \$10 to the other person, as in the standard game, or they were allowed to take up to \$2 (in \$1 increments) from the other person's show-up fee. These types of taking games have been studied by Bardsley (2008) and List (2007) who find both that a significant number of dictators take when those options are available, but also that the introduction of the additional taking options causes a downward shift in the distribution of positive amounts shared. That is, when additional taking options are introduced, many people select these options and their

[24] These and all subsequent numerical examples are obtained using the values from Table 1 and Figures 1 through 3, and incorporating them into the model in equation 1.

presence also causes people who share positive amounts in the standard variant to decrease the amount they share.

Figure 3 presents the mean ratings of social appropriateness for each action from the standard dictator game (from Table 1) and the take-2 variant. On the right side of the graph, when the dictator gives $4 or more to the recipient, the ratings are generally similar. Among these allocations, only the difference in ratings for transfers between $4 and $6 is marginally statistically significant (z = 1.69, p = 0.09), and the remaining comparisons are all statistically insignificant.

On the left side of the graph, however, the ratings differ substantially. For any amount shared with the recipient between $0 and $3, the action is significantly more socially appropriate in the take-2 variant than in the standard dictator game (z > 2.73, p < 0.01, in all four comparisons using a rank-sum test for every comparison). That is, giving small amounts to the recipient is more socially appropriate when one could have taken money instead. For example, while sharing nothing with the recipient is rated as very socially inappropriate in the standard treatment (-0.86), the same action is rated much less harshly in the take-2 variant (-0.45).

The differences in Figure 3 explain why people who share positive amounts in the standard dictator game may share less when the taking options are introduced. For example, consider again an individual for whom $V(\$10) = kV(\$5)$, and remember that this individual will prefer to share $5 over sharing nothing in the standard dictator game if $\gamma_i > (k-1)V(\$5)/1.76$. However, in the take-2 variant, the same individual will prefer to share nothing over sharing $5 as long as $(k-1)V(\$5)/1.28 > \gamma_i$.[25] Thus, for a range of concern with social norms, the same

[25] The denominator is the difference in ratings, for the take-2 variant, between keeping $5 (appropriateness rating of 0.83) and keeping $10 (-0.45).

individual will share less in the take-2 environment, where sharing little is more socially appropriate, than in the standard dictator game.

Moreover, the same subject might also be willing to take money from the recipient in the take-2 variant. For example, in Figure 3 the ratings for sharing nothing in the standard game ($10, $0) and taking $1 in the take-2 variant ($11, -$1) are almost identical. Therefore, the take-2 variant presents dictators with a way to obtain higher wealth than by sharing nothing in the dictator game, while sacrificing little in terms of disutility from violating social norms.

Therefore, consistent with List's and Bardsley's data, we find a likely two-fold effect of introducing the taking options. First, these additional options make keeping all or most of the money less socially inappropriate, thus making these actions more attractive than in the baseline. Second, these additional taking options also give dictators an opportunity to earn higher payoffs than in the baseline ($11 or $12), while sacrificing little in terms of norm-based disutility.

V.C. Information acquisition in binary dictator games

Finally, in Experiment 1 we also described two variants of a binary dictator game used by Dana, Weber and Kuang (2007). The Dana et al. experiment found that when participants (dictators) could remain ignorant about the consequences of their action for the other person, they often chose to do so in order to behave self-interestedly.[26]

In the baseline condition of their experiment, subjects chose between a ($6,$1) option, labeled "A", and one with payoffs ($5,$5), labeled "B." A majority of subjects in the role of dictator (74 percent) chose the equitable option B.

In a "hidden information" treatment, dictators were told their own payoffs ($6 for choosing A and $5 for choosing B), but not those of the other person. They were told that the

[26] This result was replicated by Munyan (2005) and Larson (2005).

payoffs for the recipient were $5 and $1, but not which payoff was associated with which action choice. Therefore, the actual payoffs could be either A:($6,$1) and B:($5,$5), as in the baseline condition, or A:($6,$5) and B:($5,$1), in which case choosing A maximized the dictators payoffs, minimized inequality, and also yielded the highest joint payoffs. The actual payoffs had been determined by a coin flip. Subjects could select to reveal the true payoffs by clicking a button, or could choose to make a choice without finding out the other person's payoffs.

In this hidden information treatment, significantly fewer subjects chose the fair option B, even though it was always at least as attractive as in the baseline condition (and often more attractive, when the payoffs were flipped). For example, in the case where the underlying payoffs were identical to those in the baseline, only 37 percent of cases resulted in a choice of B ($5,$5). Almost half of subjects (44 percent) chose not to acquire the information about the other player, even though it was costless to do so, and almost all of these subjects chose A.

To measure the social appropriateness of key actions in this experiment, we presented subjects in Experiment 1 with both the baseline and hidden information variants. For the baseline, they rated the appropriateness of choosing action "A" (yielding payoffs of ($6,$1)) or "B" (yielding payoffs of ($5,$5)). The action labels and payoffs were described similarly to the original Dana et al. experiment.

For the hidden information variant, the two possible sets of payoffs were described, as well as the dictator's opportunity to acquire the hidden payoff information. Subjects then rated four possible actions:

- not acquiring the payoff information and choosing A ($6,$?);
- not acquiring the payoff information and choosing B ($5,$?);

- acquiring the payoff information, observing the payoffs to be (($6,$1),($5,$5)), and choosing A;
- and acquiring the payoff information, observing the payoffs to be (($6,$1),($5,$5)), and choosing B.

The latter two were intended to measure whether, conditional upon acquiring information, social norms would be similar to those in the baseline.

Figure 4 presents the mean ratings for each of these possible action choices in the two treatments. In the baseline, the mean rating for choosing the fair option B (0.96, with 96 percent of subjects selecting "very appropriate") is considerably higher than the mean rating for choosing A (-0.66, with 58 percent of subjects selecting "very inappropriate" and another 35 percent selecting "somewhat inappropriate"). In the hidden information variant, the mean ratings in the case where the dictator acquires the information and finds the payoffs to be the same as in the baseline (B: 0.92; A: -0.67) are almost identical to the mean ratings in the baseline (B:0.96, A:-0.66), and the distributions of responses are almost identical.

However, in the hidden information variant, choosing not to acquire information and selecting A produces almost identical mean social appropriateness ratings (0.19) as choosing not to acquire the information and selecting B (0.19). In both cases, there are few responses of "very socially inappropriate" (about 10 percent), with the remaining responses evenly distributed among the other three options (about 30 percent for each). Thus, choosing not to acquire information, and then selecting the payoff-maximizing option, is not generally perceived as socially inappropriate to the degree that choosing selfishly (A) is when one knows the payoffs.

Going back to our model, suppose that an individual places k times as much value on receiving a payoff of $6 as on receiving a payoff of $5 (i.e., $V(\$6) = kV(\$5)$). In the baseline

treatment, such an individual will strictly prefer B to A as long as $(k-1)V(\$5)/1.62 < \gamma_i$. However, since the two ratings for the hidden information case are virtually identical, all individuals will select A conditional upon not acquiring the payoff information. Moreover, a subject will strictly prefer to not acquire the payoff information and select A, over choosing B in the baseline, as long as $(k-1)V(\$5)/0.77 > \gamma_i$. Therefore, for a range of concerns with social norms specified by the two inequalities, individuals will choose fairly (B) in the baseline condition, but would prefer to not acquire the payoff information and instead choose selfishly (A) under strategic ignorance.

VII. Conclusion

Our work makes three contributions to the study of social behavior in economics. First, we present a method for eliciting ratings over social norms using incentive compatible coordination games, and thus rely on the important property of social norms that they are jointly recognized among members of a population, rather than privately held. These simple coordination games, in which respondents rate not according to their own beliefs of the social appropriateness of actions but instead according to what they think others will respond, are a good way of measuring the socially-held and jointly-recognized perceptions regarding appropriate behavior that constitute social norms.

Our second contribution is to demonstrate that the combination of the social appropriateness ratings we elicit with a simple model that incorporates concern for norm-compliance with utility for money predicts changes across several variants of the dictator game. Thus, in addition to showing how we can measure social norms, we also demonstrate that the result of such measurement can be helpful for predicting and interpreting behavior and for

developing a model of social norm compliance. The key novel property in our model is that people care not just about monetary payoffs, but also about the relative "social appropriateness" (to other actions available to the decision maker) of any action they take.

Our third major contribution is to provide an explanation for seemingly anomalous behavior in variants of the dictator game. Both in our experiment and in related previous experiments, behavior exhibits a high degree of sensitivity to contextual features that should be largely irrelevant if people care about payoffs alone. But the changes in behavior are largely consistent with the predictions of our model and elicited social appropriateness ratings. Thus, by identifying social norms we are able to obtain a large degree of explanatory power, including *a priori* predictions, regarding how behavior changes in response to changes in the context of a choice environment.

We also demonstrate that social norms are most useful when they are viewed not as a single behavior that one should do (e.g., "share wealth equally"). Instead, we find that much of the explanatory power of social norms comes from identifying them as degrees of appropriateness over a range of possible actions available to a decision maker. Across our elicitation of norms in different variants in the dictator game, dividing money equally is always the most socially appropriate thing to do (cf. Andreoni and Bernheim 2008). However, what differs across the variants is the relative appropriateness of other actions, and our analysis suggests that this distinction (between the peak of social acceptability and the degree to which appropriateness varies across all actions) is an important way in which norms influence behavior.

Finally, we should address the extent to which what we measure using our coordination games is actually social norms. In another paper (with Rachel Croson), we applied our elicitation instrument to measure norms of tipping and punctuality, how they vary by nationality, and how

individuals can recognize distinctions in norms across populations. In the study, we also elicited perceptions of the social appropriateness of several different actions in two scenarios. But in this case we used scenarios in which we knew the norm (tipping and punctuality) and we varied the population with which participants played the coordination game (foreign students in the US playing with other students in the US or playing with people from their home country). Thus, we could compare variation in the elicited norms with variation in social norms across countries that we validated from other sources. And we could also verify whether foreign-born respondents playing the coordination game recognized that social norms differed when the reference group changed. In both cases, we find that the answer is yes. Foreign-born students provide different appropriateness ratings when matching responses with US students than with people from their home country. And the variation in ratings of social appropriateness mirror externally-validated differences in social norms. These findings are important, as they indicate that the responses we obtained track real social norms, and that respondents know that they are attempting to match socially-identified perceptions that vary on the social group.

Of course, this work represents a first step, and there remains much to be done in developing a comprehensive model of social norms. We have only demonstrated one kind of environment in which we can elicit social norms and use them to predict and interpret behavior. A comprehensive model of social norms will involve, as an example, understanding how social norms come to be jointly recognized and how they vary from reference group to reference group. We believe that the norm elicitation technique will prove useful for developing just such an understanding.

References:

Andreoni, J. and Bernheim, D. (2008). "Social image and the 50-50 Norm: A theoretical and Expeirmental Analysis of Audience Effects", *working paper.*

Andreoni, J. and Miller, J. (2003)."Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism", *Econometrica*, 70(2):737-753.

Akerlof, G. (1980). "A Theory of Social Customs, of Which Unemployment May be One Consequence," *Quarterly Journal of Economics*.

Akerlof, G. (2006). "The Missing Motivation in Macroeconomics," *Presidential Address at the American Economics Association*.

Bardsley, N. (2008). "Dictator Game Giving: Altruism or Artefact" Experimental Economics; 11(2): 122 - 133..

Berg, J., Dickhaut, J., McCabe K. (1995). "Trust, Reciprocity, and Social History", Games and Economic Behavior, 10:122-142.

Bernheim, B. D. (1994). "A Theory of Conformity". Journal of Political Economy 102(5): 841-77.

Bettenhausen, K. and Murnighan, J. (1991). "The Development of an Intragroup Norm and the Effects of Interpersonal and Structural Challenges." Administrative Science Quarterly **36**: 20-35.

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

Bicchieri, C. and Xiao, E. (forthcoming) "Do The Right Thing: But Only If Others Do So", Journal of Behavioral Decision Making.

Bohnet, Iris and Frey, Bruno S. (1999), "Social Distance and other-regarding Behavior in Dictator Games: Comment," *The American Economic Review*, Vol. 89(1), pp. 335-339.

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

Bowles, S. (1998). "Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions", *Journal of Economic Literature* 36: 75-111.

Camerer, C. and Fehr, E. (2003). "Measuring Social Norms and Preferences Using Experimental Games: A Guide for Social Scientists". Foundations of Human Sociality -- Experimental and Ethnographic Evidence from 15 Small-Scale Societies. Ed. Henrich, Boyd, Bowles, Camerer, Fehr, Gintis and McElreath.

Cialdini, R., Reno, R. and Kallgren, C. (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): 1015-1026.

Cole, H., Mailath G. and Postlewaite, A. (1992) "Social Norms, Savings Behavior and Growth," *Journal of Political Economy*, 100: 1092-1125.

Coleman, J. (1990), Foundations of Social Theory, Belknap Press of Harvard Univeristy Press (Cambridge Mass.).

Conlin, M., Lynn, M. and O'Donoghue, T. (2003). "The Norm of Restaurant Tipping." Journal of Economic Behavior and Organization **52**(3): 297-321.

Cox, J., Friedman, D. and Gjerstad, S. (2007). "A Tractable Model of Reciprocity and Fairness", Games and Economic Behavior, 59: 17-45.

Dana, J., Cain, D. and Dawes, R. (2006). "What You Don't Know Won't Hurt Me: Costly (but quiet) Exit in Dictator Games." Organizational Behavior and Human Decision Processes **100**(2): 1993-201.

Dana, J., Weber, R. and Kuang, J. X. (2007). "Exploiting Moral Wriggle Room: Behavior Inconsistent with a Preference for Fair Outcomes." Economic Theory, 33(1):67-80.

Deutsch, M. and Gerard, H. B. (1955). "A Study Of Normative And Informational Social Influences Upon Individual Judgment." The Journal of Abnormal and Social Psychology 51(3): 629-636.

Eichenberger, R. and Oberholzer-Gee, F. (1998). "Rational Moralists: The Role of Fairness in Democratic Economic Politics." Public Choice **94**(1-2): 191-210.

Elster, J. (1989). The Cement of Society: a Study of Social Order. Studies in Rationality and Social Change, Cambridge University Press.

Fehr, E. and Fischbacher, U. (2004). "Third-Party Punishment and Social Norms." Evolution and human behavior **25**(2): 63 (25 pages).

Fehr, E. and Gachter, S. (2000). "Fairness and Retaliation: The Economics of Reciprocity." Journal of Economic Perspectives.

Fehr, E. and Schmidt, K. (1999). "A Theory of Fairness, Competition, and Cooperation." The Quarterly Journal of Economics **114**(3): 817-868.

Fisher, P. and Huddart, S. (forthcoming)."Optimal Contracting With Endogenous Social Norms", *American Economic Review*.

Forsythe, R., Horowitz, J., Savin, N. and Sefton, M. (1994). "Fairness in Simple Bargaining Experiments." Games and Economic Behavior **6**(3): 347-369.

Hoffman, E., McCabe K., Shachat, K., Smith, V. (1994). "Preferences, Property Rights, and Anonymity in Bargaining Games", Games and Economic Behavior,7:346-380.

Kandori, M. (1992). “Social Norms and Community Enforcement,” *Review of Economic Studies*, 59: 62-80.

Krupka, E., Weber, R., and Croson, R.(2008). “When in Rome: Identifying Social Norms as a Group Phenomenon.” *Working Paper.*

Krupka, E. and Weber, R. (2007). The Focusing and Informational Effects of Norms on Pro-Social Behavior. IZA Working Paper 3169.Institute for the Study of Labor.

Larson, T. (2005). "The Use of Strategic Ignorance in Dictator Games When Payoffs Are Not Transparent: Computer vs. Paper Testing Techniques". Department of Economics Master’s Thesis, Emory University (2005)

Lazear, E., Malmendier, U. and Weber, R. (2006). "Sorting in Experiments with Application to Social Preferences." *NBER Working Paper* 12041.

Levitt, S. and List, J. (2007). "What Do Laboratory Experiments Measuring Social Preferences Tell Us About the Real World?" *Journal of Economic Perspectives*, 21(2): 153-174.

List, J. (2007). “On the Interpretation of Giving in Dictator Games,” *Journal of Political Economy*, 115(3): 482-494.

Mehta, J., Starmer, C. and Sugden, R. (1994). "The Nature of Salience: An Experimental Investigation of Pure Coordination Games." American Economic Review **84**(3): 658-73.

Merton, R.K. (1957) *Social Theory and Social Structure*. Free Press (Glencoe, IL).
Ross, L. and Ward, A. (1996). Naive Realism in Everyday Life: Implications for Social Conflict and Misunderstanding. Values and Knowledge. Ed. T. Brown, E. Reed and E. Turiel. Hillsdale, NJ, Erlbaum**:** 103 -135.

Munyan, L (2005). "Patterns of Information Avoidance in Binary Choice Dictator Games", Working paper.

Nee, V. (1998). "Norms and Networks in Economic and Organizational Performance", *American Economic Review* 88: 85-89.

Nyborg, K. and M. Rege, M. (2003) On "Social Norms: The Evolution of Considerate Smoking Behavior", *Journal of Economic Behavior and Organization*, 52, 323- 340.

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

Schelling, T. C. (1960). The Strategy of Conflict. Cambridge, MA, Harvard University Press.

Sugden, R. (1995). "A Theory of Focal Points", *The Economic Journal*, 105(430):533-550.

Sherif, M. (1936). The Psychology of Social Norms. Harper and Row, New York.

Swope, K., Cadigan, J., Schmitt, P., and Shupp, R. (2008). "Social Position and Distributive Justice", *Southern Economic Journal*, 74(3): 811-818.

Young, P. (2008). Social Norms. The New Palgrave Dictionary of Economics, Second Edition. Ed. S. Durlauf and L. Blume. London, Macmillan.

Young, P. (1998). “Social Norms and Economic Welfare”, *European Economic Review*, 42: 821-830.

Table 1. Elicited norms for bully vs. standard environments (Experiment 1)

Action choice (final wealth)	Standard (n = 64) (Initial wealth: $10, $0)					Bully (n = 51) (Initial wealth: $5, $5)					*rank-sum test (z)*
	Mean	- -	-	+	+ +	**Mean**	- -	-	+	+ +	
$10, $0	**-0.86**	86%	9%	3%	2%	**-0.91**	90%	8%	0%	2%	*0.71*
$9, $1	**-0.71**	64%	31%	2%	3%	**-0.84**	82%	14%	2%	2%	*2.09***
$8, $2	**-0.47**	36%	52%	9%	3%	**-0.66**	57%	37%	4%	2%	*2.29***
$7, $3	**-0.22**	11%	64%	22%	3%	**-0.37**	25%	59%	12%	4%	*1.98***
$6, $4	**0.15**	5%	23%	66%	6%	**-0.05**	8%	49%	35%	8%	*2.58****
$5, $5	**0.90**	0%	2%	13%	86%	**0.96**	0%	0%	6%	94%	*1.44*
$4, $6	**0.54**	0%	6%	56%	38%	**0.50**	6%	10%	37%	47%	*0.27*
$3, $7	**0.33**	2%	28%	39%	31%	**0.32**	8%	20%	39%	33%	*0.11*
$2, $8	**0.22**	6%	39%	20%	34%	**0.19**	16%	25%	24%	35%	*0.13*
$1, $9	**0.09**	20%	28%	19%	33%	**0.06**	29%	16%	22%	33%	*0.25*
$0, $10	**0.06**	30%	16%	20%	34%	**-0.01**	37%	12%	16%	35%	*0.41*

* - p < 0.1, ** - p < 0.05, *** - p < 0. 01; all two-tailed
Responses are: “very socially inappropriate” (- -), “somewhat socially inappropriate” (-), “somewhat socially appropriate” (+), “very socially appropriate” (+ +); modal response shaded

Table 2. Statistical tests (ordered logit) of behavior across treatments (Experiment 2)

	(1)	(2)	(3)
Dependent variable:	Amount allocated to recipient	Binary (Share = $5)	Binary (Share = $0)
Bully	0.492* (0.286)	1.376*** (0.372)	0.491* (0.300)
Class size	-0.013*** (0.005)	-0.019*** (0.005)	-0.003 (0.005)
Constant		0.949 (0.841)	0.069 (0.708)
N	91	82	58
Model:	*Ordered logistic regression*	*Logistic regression*	*Logistic regression*
Data:	*All data*	*Subjects who allocated less than $6 to recipient*	*Subjects who allocated less than $5 to recipient*

* - p < 0.1, ** - p < 0.05, *** - p < 0. 01; all two-tailed
Standard errors clustered by session

Figure 1. Distributions of amounts shared by treatment (Experiment 2)

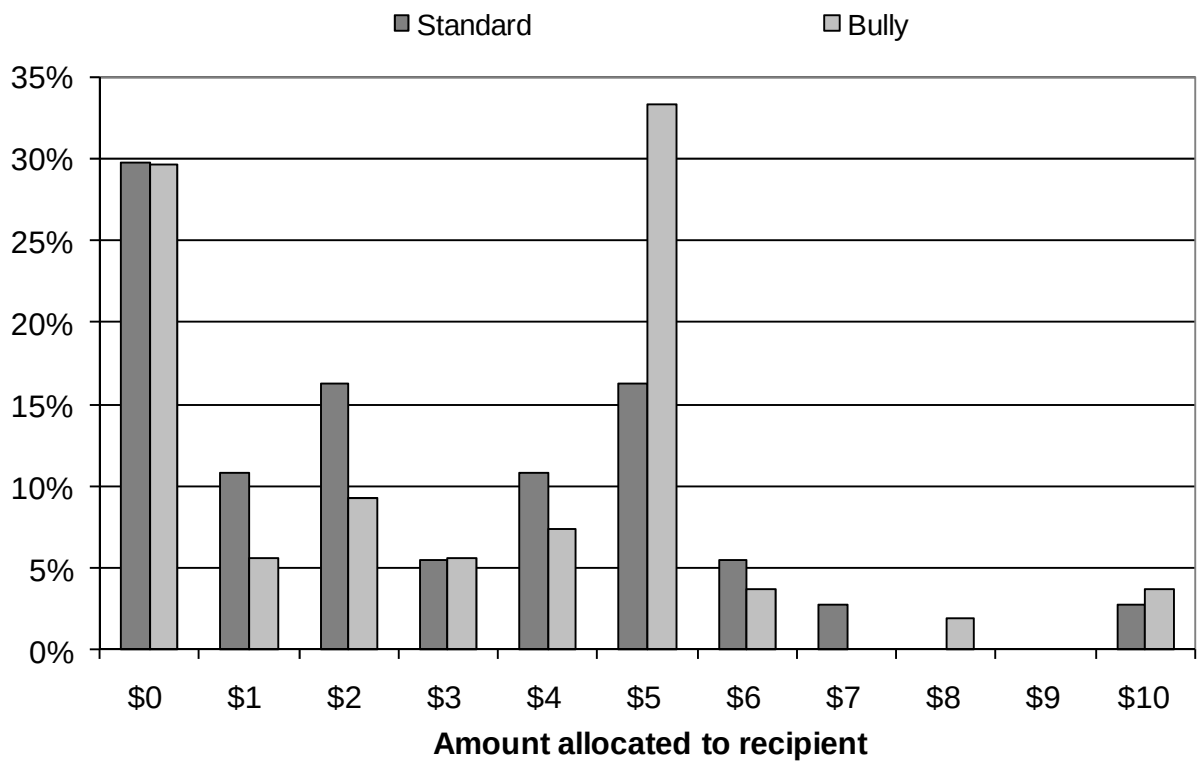


Figure 2. Mean ratings of social appropriateness (standard vs. sorting variant)

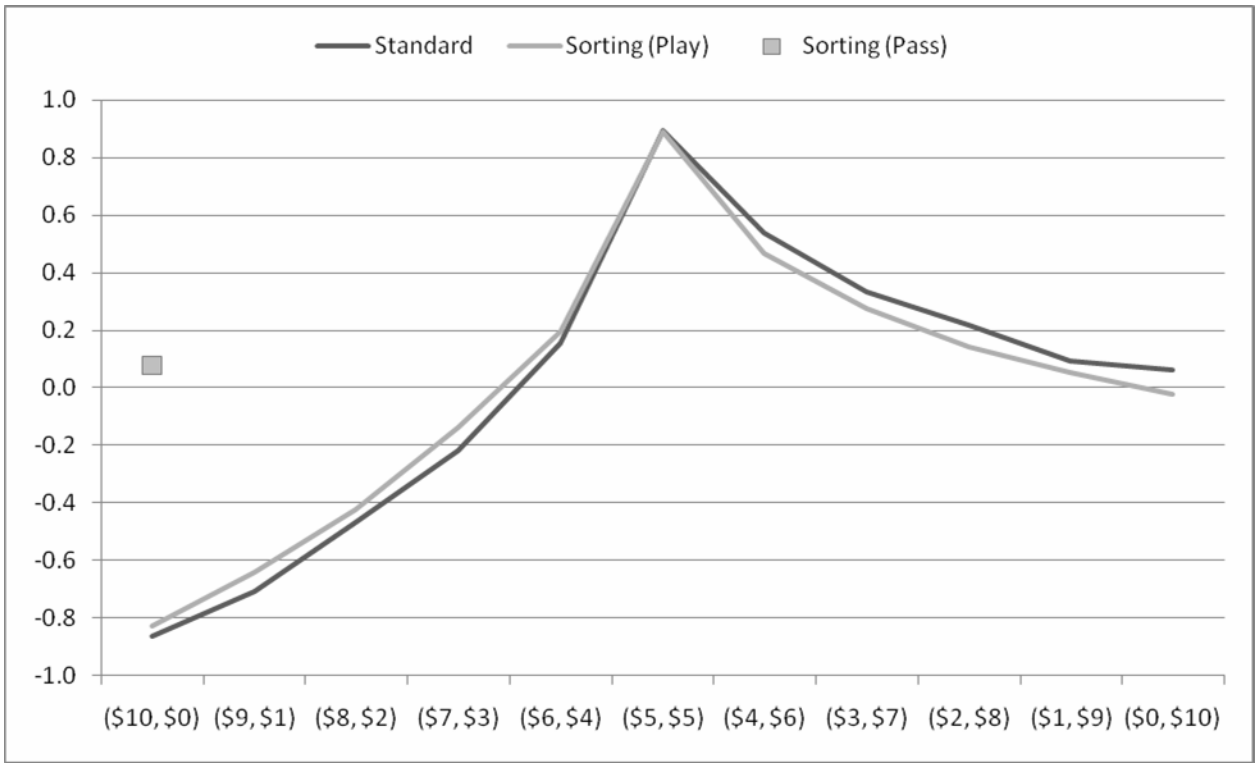


Figure 3. Mean ratings of social appropriateness (standard vs. take-2 variant)

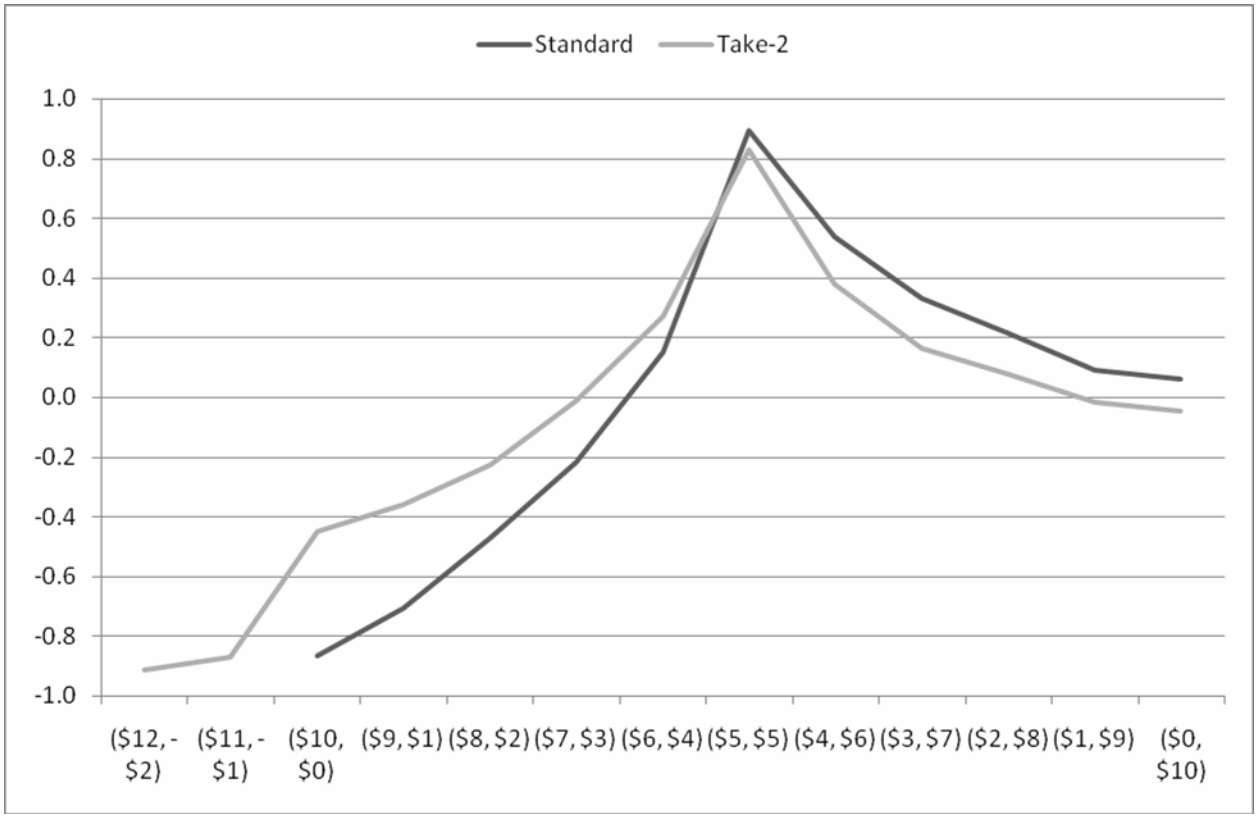
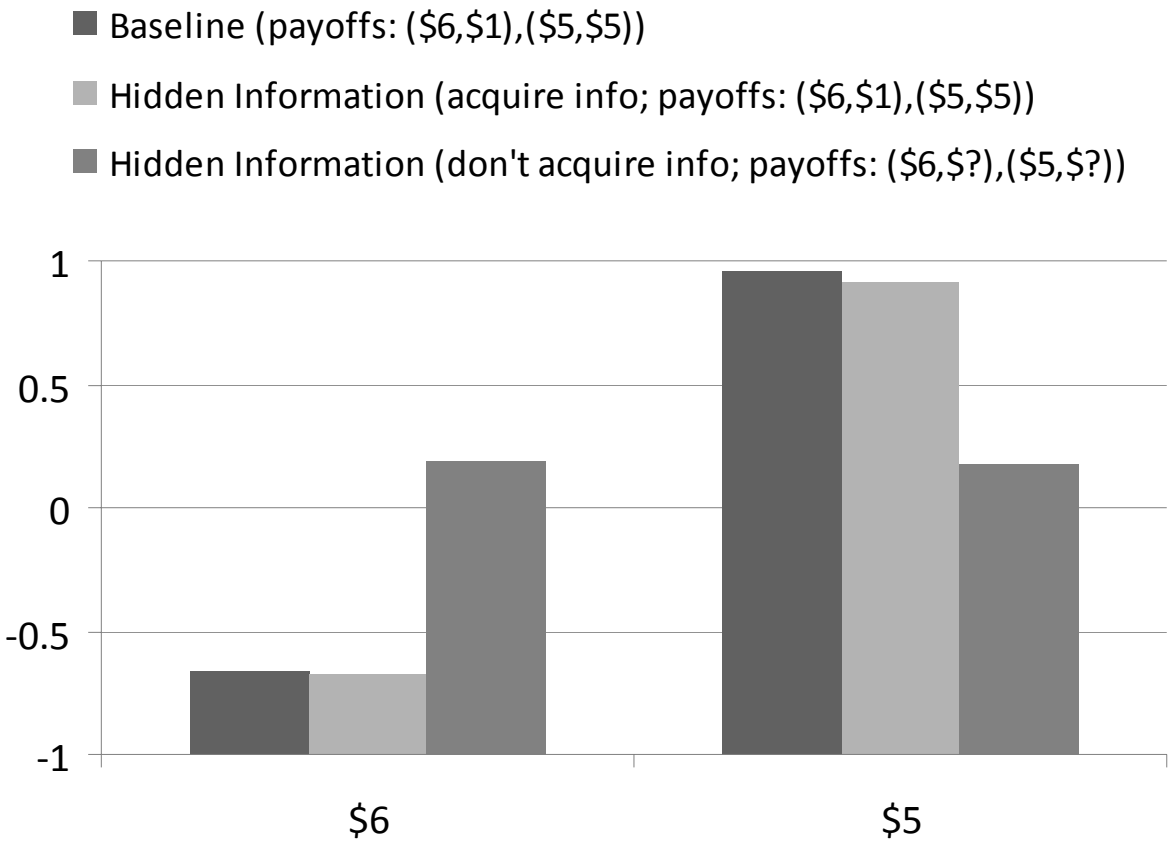


Figure 4. Mean ratings of social appropriateness (binary baseline and hidden information variants)



Appendix 1
(Instructions for Experiment 1)

Initial Instructions

(Experimenter read aloud only)

This is a study on decision making. For your participation, you will be paid a participation fee of $7. In addition, you may receive some additional money based on your choices and the choices of others during the experiment.

If you have any questions during the study, please raise your hand and wait for an experimenter to come to you. Please do not talk, exclaim, or try to communicate with other participants during the experiment. Participants intentionally violating the rules may be asked to leave the experiment and may not be paid.

Please take an envelope from the experimenter. Inside the envelope is a card with a number, from 1 to 20. This number is your Participant ID for the experiment. Please hold on to this card for the duration of the experiment. We will ask you to write this number on each set of sheets you receive.

Appendix 1
(Instructions for Experiment 1)

Participant ID: _____
(Please enter the number from your card)

Instructions, Part I

Please write your participant ID in the space provided above.

On the following pages, you will read descriptions of a series of situations. These descriptions correspond to situations in which one person, "Individual A," must make a decision. For each situation, you will be given a description of the decision faced by Individual A. This description will include several possible choices available to Individual A.

After you read the description of the decision, you will be asked to evaluate the different possible choices available to Individual A and to decide, for each of the possible actions, whether taking that action would be "*socially appropriate*" and "*consistent with moral or proper social behavior*" or "*socially inappropriate*" and "*inconsistent with moral or proper social behavior*." By socially appropriate, we mean behavior that most people agree is the "correct" or "ethical" thing to do. Another way to think about what we mean is that if Individual A were to select a socially inappropriate choice, then someone else might be angry at Individual A for doing so.

In each of your responses, we would like you to answer as truthfully as possible, based on your opinions of what constitutes socially appropriate or socially inappropriate behavior.

To give you an idea of how the experiment will proceed, we will go through an example and show you how you will indicate your responses. On the next page you will see an example of a situation.

Appendix 1
(Instructions for Experiment 1)

Example Situation

Individual A is at a local coffee shop near campus. While there, Individual A notices that someone has left a wallet at one of the tables. Individual A must decide what to do. Individual A has four possible choices: take the wallet, ask others nearby if the wallet belongs to them, leave the wallet where it is, or give the wallet to the shop manager. Individual A can choose only one of these four options.

The table below presents a list of the possible choices available to Individual A. For each of the choices, please indicate whether you believe choosing that option is very socially inappropriate, somewhat socially inappropriate, somewhat socially appropriate, or very socially appropriate. To indicate your response, please place an "x" inside one box for each row.

Individual A's choice	*Very Socially Inappropriate*	*Somewhat socially inappropriate*	*Somewhat socially appropriate*	*Very socially appropriate*
Take the wallet				
Ask others nearby if the wallet belongs to them				
Leave the wallet where it is				
Give the wallet to the shop manager				

Please make sure that you have placed one "x" in each row.

If this were one of the situations for this study, you would consider each of the possible choices above and, for that choice, indicate the extent to which you believe taking that action would be "*socially appropriate*" and "*consistent with moral or proper social behavior*" or "*socially inappropriate*" and "*inconsistent with moral or proper social behavior"*. Recall that by socially appropriate we mean behavior that most people agree is the "correct" or "ethical" thing to do.

Continue example on the next page

Appendix 1
(Instructions for Experiment 1)

For example, suppose you thought that taking the wallet was *very socially inappropriate*, asking others nearby if the wallet belongs to them was *somewhat socially appropriate*, leaving the wallet where it is was *somewhat socially inappropriate*, and giving the wallet to the shop manager was *very socially appropriate*. Then you would indicate your responses as follows:

Individual A's choice	*Very Socially Inappropriate*	*Somewhat socially inappropriate*	*Somewhat socially appropriate*	*Very socially appropriate*
Take the wallet	X			
Ask others nearby if the wallet belongs to them			X	
Leave the wallet where it is		X		
Give the wallet to the shop manager				X

Are there any questions about this example situation or about how to indicate your responses? On the following pages, there are four situations, all dealing with decisions that "Individual A," a participant in an experiment, might have to make.

For each situation, you will receive a sheet, with a table on which to indicate your responses. For each situation, the experimenter will read a description of the situation. You will then indicate whether each possible choice available to Individual A is socially appropriate or socially inappropriate.

At the end of the experiment today, we will select one of the five situations, by randomly-drawing a number from 1 to 5. For this situation, we will also randomly select one of the possible choices that Individual A could make. Thus, we will select both a situation and one possible choice at random. For the choice selected, we will determine which response was selected by the most people here today. If you give the same response as that most frequently given by other people, then you will receive an additional $5. This amount will be paid to you, in cash, at the conclusion of the experiment. For instance, if we were to select the example situation above and the possible choice "Leave the wallet where it is," and if your response had been "somewhat socially inappropriate," then you would receive $5, in addition to the $7 participation fee, if this was the response selected by most other people in today's session. Otherwise you would receive only the $7 participation fee.

If you have any questions from this point on, please raise your hand and wait for the experimenter to come to you.

Please turn the page to begin once the experimenter asks you to do so.

Appendix 1
(Instructions for Experiment 1)

Situation 1[27]

Suppose that Individual A is randomly paired with another person, Individual B. The pairing is anonymous, meaning that neither individual will ever know the identity of the other individual with whom he or she is paired. In the experiment, Individual A will make a choice, the experimenter will record this choice, and then both individuals will be informed of the choice and paid money based on the choice made by Individual A, as well as a small participation fee. Suppose that neither individual will receive any other money for participating in the experiment.

In each pair, Individual A will receive $10. Individual A will then have the opportunity to give any amount of his or her $10 to Individual B. That is, Individual A can give any of the $10 he or she receives to Individual B. For instance, Individual A may decide to give $0 to Individual B and keep $10 for him or herself. Or Individual A may decide to give $10 to Individual B and keep $0 for him or herself. Individual A may also choose to give any other amount between $0 and $10 to Individual B. This choice will determine how much money each will receive, privately and in cash, at the end of the experiment.

The table below presents a list of the possible choices available to Individual A. For each of the choices, please indicate whether you believe choosing that option is very socially inappropriate, somewhat socially inappropriate, somewhat socially appropriate, or very socially appropriate. To indicate your response, please place an "x" inside one box for each row.

Individual A's choice	*Very Socially Inappropriate*	*Somewhat socially inappropriate*	*Somewhat socially appropriate*	*Very socially appropriate*
Give $0 to Participant B (Participant A gets $10, Participant B gets $0)				
Give $1 to Participant B (Participant A gets $9, Participant B gets $1)				
Give $2 to Participant B (Participant A gets $8, Participant B gets $2)				
Give $3 to Participant B (Participant A gets $7, Participant B gets $3)				
Give $4 to Participant B (Participant A gets $6, Participant B gets $4)				
Give $5 to Participant B (Participant A gets $5, Participant B gets $5)				
Give $6 to Participant B (Participant A gets $4, Participant B gets $6)				
Give $7 to Participant B (Participant A gets $3, Participant B gets $7)				
Give $8 to Participant B (Participant A gets $2, Participant B gets $8)				
Give $9 to Participant B (Participant A gets $1, Participant B gets $9)				
Give $10 to Participant B (Participant A gets $0, Participant B gets $10)				

Please wait to turn the page until the experimenter asks you to do so. If you have any questions, please raise your hand and wait for the experimenter.

[27] This is the Standard game. This or the Bully game was always first.

Appendix 1
(Instructions for Experiment 1)

Situation 1[28]

Suppose that Individual A is randomly paired with another person, Individual B. The pairing is anonymous, meaning that neither participant will ever know the identity of the other participant with whom he or she is paired. In the experiment, Individual A will make a choice, the experimenter will record this choice, and then both participants will be informed of the choice and will be paid money based on the choice made by Individual A, as well as a small participation fee. Suppose that neither participant will receive any other money for participating in the experiment.

In each pair, Individuals A and B will each receive $5. Individual A will then have the opportunity to give any amount of his or her $5 to Individual B or to take any amount of the $5 given to Individual B for him or herself. That is, Individual A can give any of the $5 he or she receives to Individual B or can take any of the $5 Individual B receives for him or herself. For instance, Individual A may decide to not give any money and to take all $5 from Individual B. Or Individual A may decide to give all $5 to Individual B and not take any money. Individual A may also choose to give or take any amount between $0 and $5 to or from Individual B. This choice will determine how much money each participant will receive, privately and in cash, at the end of the experiment.
The table below presents a list of the possible choices available to Individual A. For each of the choices, please indicate whether you believe choosing that option is very socially inappropriate, somewhat socially inappropriate, somewhat socially appropriate, or very socially appropriate. To indicate your response, please place an "x" inside one box for each row.

Individual A's choice	*Very Socially Inappropriate*	*Somewhat socially inappropriate*	*Somewhat socially appropriate*	*Very socially appropriate*
Take $5 from Individual B (Individual A gets $10, Individual B gets $0)				
Take $4 from Individual B (Individual A gets $9, Individual B gets $1)				
Take $3 from Individual B (Individual A gets $8, Individual B gets $2)				
Take $2 from Individual B (Individual A gets $7, Individual B gets $3)				
Take $1 from Individual B (Individual A gets $6, Individual B gets $4)				
Give $0/Take $0 to/from Individual B (Individual A gets $5, Individual B gets $5)				
Give $1 to Individual B (Individual A gets $4, Individual B gets $6)				
Give $2 to Individual B (Individual A gets $3, Individual B gets $7)				
Give $3 to Individual B (Individual A gets $2, Individual B gets $8)				
Give $4 to Individual B (Individual A gets $1, Individual B gets $9)				
Give $5 to Individual B (Individual A gets $0, Individual B gets $10)				

Please wait to turn the page until the experimenter asks you to do so.
If you have any questions, please raise your hand and wait for the experimenter.

[28] This is the Bully game.

Appendix 1
(Instructions for Experiment 1)

Situation 2[29]

Suppose that Individual A is initially not paired with any other participant. During the experiment Individual A might be randomly paired with another person, Individual B. If the pairing occurs, then it will be anonymous, meaning that neither participant will ever know the identity of the other participant with whom he or she is paired.

In the experiment, Individual A will make a choice, the experimenter will record this choice, and Individual A will be paid money based on the choice, as well as a small participation fee. Another participant, Individual B, may or may not also be paid money based on the choice made by Individual A, as well as a small participation fee. Suppose that neither participant will receive any other money for participating in the experiment.

In each pair, Individual A will have a choice of whether to play a game or to pass on playing the game.

If Individual A decides to pass then he or she will receive $10, the two participants will not be paired, and Individual B will not find out anything about the game and will not receive any additional money.

If Individual A chooses to play the game, then he or she will be paired with Individual B. Individual A will receive $10 and will then have the opportunity to give any amount of his or her $10 to Individual B. That is, Individual A can give any of the $10 he or she receives to Individual B. For instance, Individual A may decide to give $0 to Individual B and keep $10 for him or herself. Or Individual A may decide to give $10 to Individual B and keep $0 for him or herself. Individual A may also choose to give any other amount between $0 and $10 to Individual B. In this case, Individual B will find out about the game and how much, if any, of the $10 Individual A chose to give and how much Individual A kept for him or herself.

If you have a question about this situation, please raise your hand and wait for the experimenter.

Situation 2 continued on the next page.

[29] Situation 2 – 5 were then presented in random order.

Appendix 1
(Instructions for Experiment 1)

Situation 2 (continued)

The table below presents a list of the possible choices available to Individual A. For each of the choices, please indicate whether you believe choosing that option is very socially inappropriate, somewhat socially inappropriate, somewhat socially appropriate, or very socially appropriate. To indicate your response, please place an "x" inside one box for each row.

Individual A's choice	*Very Socially Inappropriate*	*Somewhat socially inappropriate*	*Somewhat socially appropriate*	*Very socially appropriate*
Pass (Do not Play Game) (No pairing occurs, Individual A gets $10, Individual B is told nothing about game and gets $0)				
Play Game and Give $0 to Individual B (Individuals A and B are paired, Individual A gets $10, Individual B is told about game and gets $0)				
Play Game and Give $1 to Individual B (Individuals A and B are paired, Individual A gets $9, Individual B is told about game and gets $1)				
Play Game and Give $2 to Individual B (Individuals A and B are paired, Individual A gets $8, Individual B is told about game and gets $2)				
Play Game and Give $3 to Individual B (Individuals A and B are paired, Individual A gets $7, Individual B is told about game and gets $3)				
Play Game and Give $4 to Individual B (Individuals A and B are paired, Individual A gets $6, Individual B is told about game and gets $4)				
Play Game and Give $5 to Individual B (Individuals A and B are paired, Individual A gets $5, Individual B is told about game and gets $5)				
Play Game and Give $6 to Individual B (Individuals A and B are paired, Individual A gets $4, Individual B is told about game and gets $6)				
Play Game and Give $7 to Individual B (Individuals A and B are paired, Individual A gets $3, Individual B is told about game and gets $7)				
Play Game and Give $8 to Individual B (Individuals A and B are paired, Individual A gets $2, Individual B is told about game and gets $8)				
Play Game and Give $9 to Individual B (Individuals A and B are paired, Individual A gets $1, Individual B is told about game and gets $9)				
Play Game and Give $10 to Individual B (Individuals A and B are paired, Individual A gets $0, Individual B is told about game and gets $10)				

Please wait to turn the page until the experimenter asks you to do so.
If you have any questions, please raise your hand and wait for the experimenter.

Appendix 1
(Instructions for Experiment 1)

Situation 3

Suppose that Individual A is randomly paired with another person, Individual B. The pairing is anonymous, meaning that neither participant will ever know the identity of the other participant with whom he or she is paired.

In the experiment, Individual A will make a choice, the experimenter will record this choice, and then both participants will be paid money based on the choice made by Individual A, as well as a small participation fee. Suppose that neither participant will receive any other money for participating in the experiment.

In each pair, Individual A must choose between two options, labeled "X" and "Y". If Individual A chooses X he or she receives $6 and Individual B receives $1. If Individual A chooses Y he or she receives $5 and Individual B receives $5. The outcomes produced by Individual A's choices are presented in the following payoff table.

Individual A's choice	*Payoff to Individual A*	*Payoff to Individual B*
X	$6	$1
Y	$5	$5

Individual B is aware of the options faced by Individual A. Individual B will find out how much money he or she received, after Individual A makes a choice. This choice will determine how much money each participant will receive, privately and in cash, at the end of the experiment.

The table below presents a list of the possible choices available to Individual A. For each of the choices, please indicate whether you believe choosing that option is very socially inappropriate, somewhat socially inappropriate, somewhat socially appropriate, or very socially appropriate. To indicate your response, please place an "x" inside one box for each row.

Individual A's choice	*Very Socially Inappropriate*	*Somewhat socially inappropriate*	*Somewhat socially appropriate*	*Very socially appropriate*
X (Individual A gets $6, Individual B gets $1)				
Y (Individual A gets $5, Individual B gets $5)				

Please wait to turn the page until the experimenter asks you to do so.
If you have any questions, please raise your hand and wait for the experimenter.

Situation 4

Suppose that Individual A is randomly paired with another person, Individual B. The pairing is anonymous, meaning that neither participant will ever know the identity of the other participant with whom he or she is paired.

In the experiment, Individual A will make a choice, the experimenter will record this choice, and then both participants will be paid money based on the choice made by Individual A, as well as a small participation fee. Suppose that neither participant will receive any other money for participating in the experiment.

In each pair, Individual A must choose between two options, labeled "X" and "Y". If Individual A chooses X he or she receives $6. If Individual A chooses Y he or she receives $5. The consequences for Individual B depend on one of two scenarios, Scenario 1 and Scenario 2, each of which is equally likely. Individual A does not know initially which scenario he or she faces. Therefore, the outcomes produced by Individual A's choices are presented in the following payoff table.

Individual A's choice	*Payoff to Individual A*	*Payoff to Individual B*
X	$6	?
Y	$5	?

The actual outcomes for Individual B depend on which of two scenarios applies to the decision, Scenario 1 or Scenario 2. The scenario that applies was determined prior to the experiment by a coin flip, meaning that either scenario is equally likely.

- In Scenario 1: If Individual A chooses X then Individual B receives $1 (and Individual A receives $6). If Individual A chooses Y then Individual B receives $5 (and Individual A receives $5).
- In Scenario 2: If Individual A chooses X then Individual B receives $5 (and Individual A receives $6). If Individual A chooses Y then Individual B receives $1 (and Individual A receives $5).

The two possible scenarios are presented in the table below.

Scenario 1				**Scenario 2**		
Individual A's choice	*Payoff to Individual A*	*Payoff to Individual B*		*Individual A's choice*	*Payoff to Individual A*	*Payoff to Individual B*
X	$6	$1		**X**	$6	$5
Y	$5	$5		**Y**	$5	$1

Situation 4 continued on the next page.

Appendix 1
(Instructions for Experiment 1)

Situation 4 (continued)

Individual A must make the following choices.

- First, he or she must decide whether or not to find out the actual scenario. That is, Individual A can find out whether the payoffs are determined by Scenario 1 or Scenario 2. Individual B will not know whether A finds out or does not.
- Second, after deciding whether to find out the actual scenario, Individual A will make a choice of X or Y. If Individual A chooses to find out the actual scenario, then he or she will know precisely which payoffs for Individual B will be produced by his or her choice. If Individual A chooses not to find out the actual scenario, then he or she will not know precisely which payoffs for Individual B will be produced by his or her choice. Recall that in either scenario, Individual A knows that he or she receives $6 for choosing X and $5 for choosing Y. The difference between the two scenarios is based only on how much money Individual B receives for each of the two choices.

Individual B is aware of the decision faced by Individual A. Individual B will find out how much money he or she received, after Individual A decides whether or not to find out the actual scenario and makes a choice. However, Individual B will not find out whether Individual A chose to find out the actual scenario. Individual A's choice will determine how much money each participant will receive, privately and in cash, at the end of the experiment.

If you have a question about this situation, please raise your hand and wait for the experimenter.

The table below presents a list of the possible choices available to Individual A. For each of the choices, please indicate whether you believe choosing that option is very socially inappropriate, somewhat socially inappropriate, somewhat socially appropriate, or very socially appropriate. To indicate your response, please place an "x" inside one box for each row.

Individual A's choice	*Very Socially Inappropriate*	*Somewhat socially inappropriate*	*Somewhat socially appropriate*	*Very socially appropriate*
A does NOT find out the actual scenario and chooses X (Individual A gets $6, Individual B gets $1 or $5, Individual A does not know what Individual B will get)				
A does NOT find out the actual scenario and chooses Y (Individual A gets $5, Individual B gets $5 or $1, Individual A does not know what Individual B will get)				
A finds out the actual scenario is Scenario 1, chooses X (Individual A gets $6, Individual B gets $1)				
A finds out the actual scenario is Scenario 1, chooses Y (Individual A gets $5, Individual B gets $5)				

Please wait to turn the page until the experimenter asks you to do so.
If you have any questions, please raise your hand and wait for the experimenter.

Appendix 1
(Instructions for Experiment 1)

Situation 5

Suppose that Individual A is randomly paired with another person, Individual B. The pairing is anonymous, meaning that neither individual will ever know the identity of the other individual with whom he or she is paired. In the experiment, Individual A will make a choice, the experimenter will record this choice, and then both individuals will be informed of the choice and paid money based on the choice made by Individual A, as well as a small participation fee. Suppose that neither individual will receive any other money for participating in the experiment.

In each pair, Individual A will receive $10. Individual A will then have the opportunity to give any of this $10 to Individual B or to take up to $2 of Individual B's participation fee. For instance, Individual A may decide to take $2 from Individual B and keep the $10 for him or herself. Or Individual A may decide to give all $10 to Individual B and not take any money. Individual A may also choose to give any other amount between $0 and $10 to Individual B or to take any amount between $0 and $2 from Individual B. This choice will determine how much money each will receive, privately and in cash, at the end of the experiment.

The table below presents a list of the possible choices available to Individual A. For each of the choices, please indicate whether you believe choosing that option is very socially inappropriate, somewhat socially inappropriate, somewhat socially appropriate, or very socially appropriate. To indicate your response, please place an "x" inside one box for each row.

Individual A's choice	*Very Socially Inappropriate*	*Somewhat socially inappropriate*	*Somewhat socially appropriate*	*Very socially appropriate*
Take $2 from Participant B (Participant A gets $12, Participant B loses $2)				
Take $1 from Participant B (Participant A gets $11, Participant B loses $1)				
Give/Take $0 to/from Participant B (Participant A gets $10, Participant B gets $0)				
Give $1 to Participant B (Participant A gets $9, Participant B gets $1)				
Give $2 to Participant B (Participant A gets $8, Participant B gets $2)				
Give $3 to Participant B (Participant A gets $7, Participant B gets $3)				
Give $4 to Participant B (Participant A gets $6, Participant B gets $4)				
Give $5 to Participant B (Participant A gets $5, Participant B gets $5)				
Give $6 to Participant B (Participant A gets $4, Participant B gets $6)				
Give $7 to Participant B (Participant A gets $3, Participant B gets $7)				
Give $8 to Participant B (Participant A gets $2, Participant B gets $8)				
Give $9 to Participant B (Participant A gets $1, Participant B gets $9)				
Give $10 to Participant B (Participant A gets $0, Participant B gets $10)				

Once you are done, please turn the set of sheets over and wait for the experimenter to collect them. Make sure that your Participant ID is written on the front page.

Appendix 2
(Instructions for Experiment 2, Standard)

Instructions
(Read aloud to all subjects)

In this experiment you will be paired with one of the people in the other group. You will never know the identity of the person with whom you are paired and this person will never know your identity.

Everyone received a $2 participation payment in an envelope. Please put this envelope away, as it is yours to keep.

In this experiment, you will make a simple choice that will determine how much additional money both you and the person with whom you are paired will receive.

You currently have an envelope that contains 10 $1 bills. At the front of the room, I have a yellow envelope for each person in the other group, which contains no bills. Therefore, you currently have $10 and the other person has $0 dollars.

You may choose how much, if any, of your $10 to give to the other person. This is the only choice you will make in today's experiment. The other person will not make any choice. Both you and the other person will receive only the money determined by your choice, in addition to the $2 participation payment.

In a moment, you will each come to the front of the room one at a time. When you come to the front, you will receive a yellow envelope labeled "Money for other person." You will then exit the room and will privately decide how many, if any, of the $10 in your white envelope you would like to place inside the other person's yellow envelope. No one will ever know what choice you made.

Once you are done allocating the money as you would like, you will place the yellow envelope inside the box outside this room. You will then be finished with the experiment and may leave.

After everyone from the first group has left, the people in the other group will exit the room one at a time and will each receive one of the yellow envelopes at random. Any amount inside that yellow envelope will be theirs to keep, in addition to the $2 participation bonus.

If you have a question, please raise your hand and I will come to you. Once I have answered any questions, we will start.

Appendix 2
(Instructions for Experiment 2, Bully)

Instructions
(Read aloud to all subjects)

In this experiment you will be paired with one of the people in the other group. You will never know the identity of the person with whom you are paired and this person will never know your identity.

Everyone received a $2 participation payment in an envelope. Please put this envelope away, as it is yours to keep.

In this experiment, you will make a simple choice that will determine how much additional money both you and the person with whom you are paired will receive.

You currently have an envelope that contains 5 $1 bills. At the front of the room, I have a yellow envelope for each person in the other group, which also contains 5 $1 bills. Therefore, you currently have $5 and the other person has $5.

You may choose how much, if any, of your $5 to give to the person or how much, if any, of the other person's $5 to take for yourself. This is the only choice you will make in today's experiment. The other person will not make any choice. Both you and the other person will receive only the money determined by your choice, in addition to the $2 participation payment.

In a moment, you will each come to the front of the room one at a time. When you come to the front, you will receive a yellow envelope labeled "Money for other person." This envelope will contain 5 $1 bills. You will then exit the room and will privately decide how many, if any, of the $5 in your white envelope you would like to place inside the other person's yellow envelope or how much, if any, of the other person's $5 you would like to place inside your white envelope. No one will ever know precisely how much money you placed in or took from the other person's envelope.

Once you are done allocating the money as you would like, you will place the yellow envelope inside the box outside this room. You will then be finished with the experiment and may leave.

After everyone from the first group has left, the people in the other group will exit the room one at a time and will each receive one of the yellow envelopes at random. Any amount inside that yellow envelope will be theirs to keep, in addition to the $2 participation bonus.

If you have a question, please raise your hand and I will come to you. Once I have answered any questions, we will start.