Management Science

Strategic Ignorance and the Robustness of Social Preferences

Zachary Grossman

Strategic Ignorance and the Robustness of Social Preferences

Zachary Grossman
Department of Economics, University of California, Santa Barbara, Santa Barbara, California 93106,
grossman@econ.ucsb.edu

Participants in dictator games frequently avoid learning whether their choice to maximize their own earnings will help or hurt the recipient and then choose selfishly, exploiting the "moral wiggle room" provided by their ignorance. However, this is found in an environment in which the dictator must actively learn the true payoffs, so inaction means ignorance. Does this effect persist when one must actively choose either to be ignorant or to be informed or when one must actively choose to remain ignorant? In fact, whereas 45% of dictators remain ignorant when one must click to become informed, this drops to 25% when one must click in either case and to 3% when one must click to remain ignorant. Although the exploitation of moral wiggle room is not merely an artifact, it is, much like social behavior itself, subject to environmental and psychological factors that may reinforce or undermine its impact.

Data, as supplemental material, are available at http://dx.doi.org/10.1287/mnsc.2014.1989.

1. Introduction

Participants in experiments regularly give up money to help anonymous recipients. To the extent that social preferences influence economic behavior in the field, they should shape our understanding of competition, cooperation, incentives and contracting, property rights, social norms, and market failures (Fehr and Fischbacher 2002). Yet in many natural social environments, a potential giver can manipulate her own choice set, gather information about her options, or affect the information that others have about her choice. Experiments that incorporate such features into the decision environment have painted a portrait of moral behavior as fickle and flickering. For example, participants in dictator games avoid learning whether their choice to maximize their own earnings will help or hurt the recipient and then choose to maximize their own earnings (see Broberg et al. 2007; Dana et al. 2006, 2007; Lazear et al. 2012).

These experiments call into question the importance of social preferences and the prevalence of unselfish behavior in natural settings.[1] They also highlight how *strategic ignorance*, the willful avoidance of evidence about the negative social impact of one's own decisions—which has played an important role in political and corporate corruption, conflict, and even genocide—can perpetuate selfish and harmful behavior. This paper presents an experiment that sheds light on how environmental or psychological factors undermine the expression of social preferences and, thus, their impact in the field. The results show that, although the opportunity to remain ignorant of the impact of one's choice does indeed undermine prosocial behavior, requiring ignorance to be actively chosen reinforces prosocial behavior, so the exploitation of moral wiggle room is itself a delicate phenomenon.

The experiment uses the moral wiggle room game of Dana et al. (2007) (hereafter, DWK), in which a dictator facing the option to lower her own payoff is initially uncertain whether that sacrifice will help or hurt the recipient but can click a button to reveal this information. DWK demonstrate that many dictators will avoid this costless information, and as a result,

---

[1] Because they document behavior inconsistent with many "first-generation" social preference models, which largely focused on preferences over either reciprocity or the distribution of material payoffs, they also present an important challenge to the broader social preferences literature. For example, Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) consider distributional preferences, Rabin (1993) and Dufwenberg and Kirchsteiger (2004) consider

reciprocity, and Charness and Rabin (2002) and Falk and Fischbacher (2006) consider both. This has partly prompted a shift in focus to belief-dependent psychological motivations such as concern for self- or social image, guilt, shame, blame, and responsibility. See, for example, Bénabou and Tirole (2006, 2011), Ariely et al. (2009), Andreoni and Bernheim (2009), Charness and Dufwenberg (2006), and Grossman (2010).

significantly fewer dictators take the simple steps necessary to guarantee a fair outcome than do in a baseline dictator game with full information. Beyond simply behaving more selfishly when the link between actions and outcomes is obscured, DWK's participants actually seek to obscure this link, using ignorance strategically.

I adapt the game by varying the default information choice—specifically, whether the decision to reveal the recipient's payoffs is an active choice with no default or whether the default choice is to reveal or not reveal the payoffs. Although one condition reproduces DWK's findings, the purpose of this study is not to provide a replication, because a handful of other studies accomplish this under varying conditions.[2] Instead, my results show that in an environment in which strategic ignorance and behavior inconsistent with a preference for fair outcomes prevail, subtle changes to the information-choice elicitation method drastically alter behavior. Requiring an active choice or switching the default choice to "reveal" decreases the rate of dictator ignorance of the true state from 45% to 25% to 3% across the conditions. As a result, behavior in these variants of the moral wiggle room game *is* consistent with a preference for fair outcomes.

To better understand whether the ignorance is strategic in nature, I include a separate condition in which the dictator states her outcome choice using the strategy method. This condition maintains ignorance as the default information choice, but it removes its usefulness as a tool that allows the dictator to avoid expressing her choice in a moral dilemma. The high rate at which participants in this condition overcome whatever bias favors the default and actively choose to reveal suggests that ignorance is not merely a product of the default and that strategic considerations play an important role in a person's decision to avoid information.

People decline information to avoid facing an informed decision, not merely to avoid learning the state. However, having to actively choose ignorance is sufficiently internally costly in that it overcomes the strategic benefit of avoiding the informed decision, so behavior is inconsistent with baseline giving only when ignorance is the default choice. The exploitation of moral wiggle room is not merely an accident or artifact; it is, much like social behavior itself, subject to environmental and psychological factors that may reinforce or undermine its impact. In particular, social preferences can be expected to have a greater impact when information about the consequences of one's choices cannot be passively avoided.

2. Experimental Design

The experiments took place at the Experimental and Behavioral Economics Laboratory (EBEL) at the University of California, Santa Barbara (UCSB). Participants were randomly recruited from the EBEL subject pool (largely comprising UCSB students and staff) using the online system ORSEE (Online Recruitment System for Economic Experiments; Greiner 2003). Upon arriving at the experiment, participants sat at computer terminals, on which they read the instructions and entered their decisions. The interface was programmed using the z-Tree software package (Fischbacher 2007).[3]

Participants played the same binary dictator game used by DWK. They were instructed that they would be playing a simple game with one other person in the room with whom they had been randomly and anonymously matched, with both members of the group being paid according to the choice of the dictator. The dictator was referred to as "Player X" and the recipient as "Player Y." After participants read instructions describing a generic payoff table, they completed a short quiz to ensure that the task was understood. Next they were shown the actual payoffs for the experiment and any other information relevant to their particular experimental condition before taking another short quiz. Before participants were told which role they had been assigned, they were allowed to make a choice and were given 60 seconds—during which the payoff matrix or matrices were displayed on the screen—to consider their choice. Unless otherwise noted, the screen progression and layout reproduced the DWK interface as faithfully as possible. The text of the general instructions was reproduced almost verbatim, as were the condition-specific instructions in the replication conditions.[4]

The dictator always received $6 for choosing option A and $5 for choosing B, but there were two possibilities for the recipient's payoffs. In the "conflicting interests" (*CI*) version of the game, the recipient's payoffs from A and B were $1 and $5, respectively, whereas in the "aligned interests" (*AI*) version, the recipient's payoffs were flipped—$5 and $1, respectively. As in DWK, the *Baseline* condition featured a transparent dictator game featuring the CI payoffs, whereas the remaining conditions relaxed this transparency by having the dictator initially unsure of the version, having been told that each was equally likely.

While the dictators made their choices, recipients were asked to choose hypothetically between the two

---

[2] Larson and Capra (2009) use a paper-based, double-blind procedure that required an active information choice, lacking a default. They largely replicate DWK's results, albeit with a much smaller sample. Also, Feiler (2007) varies probabilities and payoffs, and Matthey and Regner (2007) produce a similar result in a social dilemma.

[3] Screenshots of the instructions as well as the decision interfaces for each condition are available in the online appendix at http://econ.ucsb.edu/~grossman/papers/strategic_ignorance_appendix.pdf.

[4] Minor differences in layout arose because the DWK experiment was programmed using a different software package.

options for each of the two versions of the game, with the exception that in the Baseline condition they were only asked about the CI version. After making their choices, participants were asked about their beliefs. All recipients were asked to state their beliefs about the percentage of dictators in their session who, in the relevant conditions, were knowingly playing the conflicting interests game and chose option B and, in the relevant conditions, the percentage of dictators who chose to reveal the payoffs. Dictators were asked to state their beliefs about the mean of the responses of the recipients in the same session.[5] Upon completion of the experiment, participants were paid privately in cash.

The five experimental conditions are described below. The first two exactly replicate the DWK experiment, whereas three additional conditions feature variations of the choice elicitation method. The Default NR, Active Choice, and Default R conditions are designed to provide insight into the role of the default choice in determining the dictator's choice of information. The Strategy Method condition highlights the dictator's preferences regarding acquiring information that cannot in any way affect her choice (because the choice was already made on a binding, contingent basis).[6]

1. *Baseline*: This condition exactly replicated the DWK baseline condition. Dictators played the "conflicting interests" game with certainty. Recipients made a hypothetical choice for the same game. The link between actions and outcomes was transparent. Then the participants answered the role-specific beliefs-elicitation question.

2. *Default NR*: This condition replicated the "hidden information" condition of DWK. Participants were presented with the two versions of the game and told that the true payoffs were equally likely and would never be revealed publicly but that the dictator could reveal them by clicking a button. In contrast to DWK's procedure, in which the reveal decision was made on the same screen as the outcome choice, the dictator was required to choose whether or not to reveal before proceeding to the screen with the outcome choice. This deviation from DWK's design served to maintain comparability across the three default conditions. The

information-choice screen featured two buttons, labeled "Proceed" and "Reveal Game," with the "Proceed" button preselected. Participants could click either one of these buttons but could also simply click "Ok" to proceed to the outcome-choice screen. Failing to click the "Reveal Game" button preserved the uncertainty, making ignorance the default choice. Players in both roles were informed that the dictator's decision of whether to reveal would be kept private. Thus, the dictator could remain ignorant of the payoffs, and the recipient would not know her information state.

3. *Active Choice*: This condition differed from the Default NR condition only in that the two buttons were proceeded with the text "Reveal Game" and were labeled "No" and "Yes," respectively, with neither preselected as a default. Thus, there was no status quo or default choice, and participants were required to actively click on one of the buttons before proceeding to the outcome-choice screen.

4. *Default R*: This condition differed from the Default NR condition only in that the default information choice was to reveal the recipients' payoffs. The two buttons were labeled "Proceed" and "Do Not Reveal Game," with the "Proceed" preselected. Participants could click either one of these buttons but could also simply click "Ok" to proceed to the outcome-choice screen. Failing to click the "Do Not Reveal Game" button resolved the payoff uncertainty, so the status quo or default choice was to reveal. Players in both roles were informed that the dictator's information choice would be kept private. Thus, the dictator could remain ignorant of the payoffs, and the recipient would not know her information state.

5. *Strategy Method*: In this condition, the dictator entered her outcome choice for each of the two payoff schemes, with the outcome determined by her choice in the realized version. Furthermore, the information choice was made on the same screen, rather than on a separate, preceding screen. Below where they made the outcome choice, dictators could click a button to have the payoffs revealed on a subsequent screen, but they were bound by their contingent outcome choices, regardless of their information choice or the payoff realization. As in the Default NR condition, participants could proceed without clicking the button, in which case the payoffs were by default not revealed. Only by actively clicking the "Reveal Game" button would players be later told the payoffs.

The precise elicitation method and interface was not shown to the recipient in any condition, so conditions 2–5 were identical from the perspective of the recipient. Thus, all differences in behavior across these conditions are due to direct effects of the elicitation-method manipulation. Each subject participated in only one condition and was not aware of the other conditions.

---

[5] Participants whose responses were within five percentage points of the correct answer were paid an additional $5. This elicitation method has previously been used by Dufwenberg and Gneezy (2000) and Charness and Dufwenberg (2006) and has the advantages of simplicity and immunity to risk aversion. Results of these decisions are not central to the questions and conclusions of this paper, but they are summarized in the online appendix.

[6] It is not obvious that choices are different here with the strategy method than with the usual direct-response method because one makes choices with the same lack of stimuli in both cases. The evidence presented in Brandts and Charness (2011) does not generally find qualitative differences between these elicitation methods.

Figure 1   The Percentage of Dictators Choosing Ignorance Drops Sharply as the Information-Choice Elicitation Changes from Default NR to Active Choice to Default R
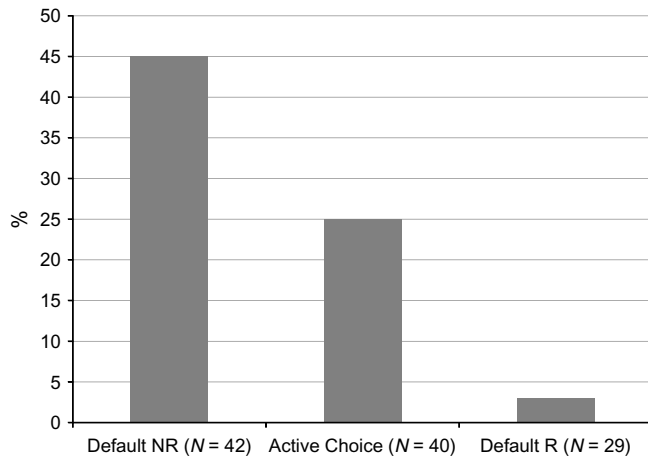
A total of 344 subjects participated across the five conditions, with exactly half (172) playing the role of dictator (Player X). On average, participants earned $10.53, including a $5 show-up fee and incentive payments for the belief elicitation. Dictators earned slightly more ($11.40) than recipients ($9.67). Sessions lasted approximately 20 minutes.

3.   Results

Figure 1 shows how the information-choice elicitation method affects the rate at which dictators choose ignorance. The percentage of subjects choosing ignorance drops sharply as the information-choice elicitation changes from Default NR to Active Choice to Default R. Of the 40 participants in the Active Choice condition, 10 (25%) chose ignorance, significantly fewer ($Z = 1.92$, $p < 0.03$) than the 19 out of 42 (45%) that did in the Default NR condition.[7] Ignorance almost disappears in the Default R condition, with only 1 of 29 participants (3%) choosing not to reveal. This low ignorance rate is significantly different from that found in both the Default NR condition ($Z = 3.85$, $p < 0.001$) and the Active Choice condition ($Z = 2.41$, $p < 0.01$). Thus, the ignorance found by DWK and replicated in the Default NR condition heavily depends on participants' ability to choose it passively.[8]

[7] A previous version of this paper reported results for a different implementation of the Default NR condition, in which the information choice and outcome choice were made on the same screen. That implementation is excluded because the Default NR condition reported herein is more directly comparable with the other default conditions. However, its results provide further support for the difference between the Default NR and Active Choice conditions, with 21 out of 39 (54%) choosing ignorance.

[8] Unless otherwise noted, all hypothesis tests are one-tailed, pooled-sample tests of the difference of proportions.

Table 1   Dictators in the Strategy Method Condition Chose Ignorance at Low Rates, Regardless of Whether Their Contingent Strategy Would Reveal the Payoff State Ex Post

	% choosing ignorance	N
Strategy Method	26	35
Unconditional (AA) choosers	29	17
Condition (BA) choosers	27	15

Table 1 displays results from the Strategy Method condition. Overall, 9 out of 35 dictators (26%) chose ignorance, significantly fewer ($Z = 1.77$, $p < 0.04$) than the 45% in the Default NR condition. Among the 17 dictators who chose A unconditionally (i.e., in both versions of the game), 5 (29%) chose ignorance. The 15 who chose B only in the CI game remained ignorant at a similar rate, with 4 (27%) choosing ignorance. These conditional choosers would learn the payoff state simply by observing their own payoff at the end of the session, but instead, they overcame the default to select "reveal" to resolve their uncertainty immediately. The high rate at which they did this suggests that the cost of overcoming the default is small relative to the desire to resolve payoff uncertainty ex post. It seems that many of the dictators who chose to remain ignorant when knowledge could affect their choice are actually curious about the final resolution.

For the sake of establishing both the faithfulness of the replication of DWK's experiment and the robustness of their results, Table 2 displays the results of the Baseline and Default NR conditions alongside the results of the equivalent conditions in DWK. Whereas only 9 out of 26 (35%) dictators in the Baseline condition chose (6, 1), in the Default NR condition, 25 of 42 (60%) chose in a manner inconsistent with a preference for the fair outcome, either choosing to remain ignorant of the recipient's payoffs or, conditional on revealing and being in the CI game, choosing (6, 1). The 25-percentage-point difference is significant at the 2% level ($Z = 2.00$). Thus, the main result of DWK is replicated.

Furthermore, the replication lines up closely with the original, differing little along several key measures, with

Table 2   In the Baseline and Default NR Conditions, the Percentage Choosing Inconsistently with a Preference for Fair Outcomes or Choosing Ignorance Is Comparable to That Found by DWK

	Replication	DWK
Baseline		
N	26	19
Inconsistent with preference for (5, 5)	35	26
Default NR		
N	39	32
Inconsistent with preference for (5, 5)	60	53
Ignorance	45	44
(5, 5) given revealed and in CI game (N)	54 (13)	75 (8)
$-maximizing given ignorant (N)	89 (19)	86 (14)

Table 3   Only in the Default NR Condition Is Behavior Significantly More Inconsistent with Fairness Preferences Than in the Baseline

	% inconsistent	N
Baseline	35	26
Default NR	60	42
Active Choice	35	40
Default R	28	29

none of the differences being statistically significant. These measures include the Baseline giving percentage (35 versus 26, $Z = 0.59$); percentage of choices inconsistent with a preference for fair outcomes (60 versus 53, $Z = 0.55$); percentage choosing to remain ignorant (45 versus 44, $Z = 0.13$); percentage choosing $(5, 5)$, conditional on revealing and facing the CI game payoffs (54 versus 75, $Z = 0.97$); and percentage of ignorant participants choosing selfishly (89 versus 86, $Z = 0.33$).[9] Thus, the replication faithfully reproduced all the basic findings of DWK.

In the Baseline condition, choosing $(6, 1)$ is not consistent with a preference for the fair outcome. In the Default NR, Active Choice, and Default R conditions, a dictator who remains ignorant of the recipient's payoffs or knowingly chooses $(6, 1)$ in the CI game is similarly inconsistent. Table 3 compares the percentage of dictators exhibiting such inconsistent behavior between the Baseline condition and the default conditions. As noted previously, behavior in the Default NR condition is not consistent with the Baseline giving rate. However, as ignorance diminishes with the changes in the default, so does the moral wiggle room that allows people to act more selfishly than in the Baseline condition. In the Active Choice condition, the percentage of choices inconsistent with a preference for fairness is 35%, exactly matching the Baseline condition, and the Default R condition features *fewer* inconsistent choices (28%) than the Baseline. Thus, it is only when ignorance can be chosen passively that people exploit moral wiggle room to obtain selfish outcomes.

4.   Discussion and Conclusion

The sharp decline in the rate at which dictators choose to remain ignorant of the recipients' payoffs across the Default NR, Active Choice, and Default R conditions shows that information acquisition decisions are significantly affected by the default information choice. Ignorance largely disappears when it cannot be

chosen passively. Although a large body of research documents status quo bias among individuals facing financial or consumer choices (e.g., Madrian and Shea 2001, Samuelson and Zeckhauser 1988), this result demonstrates that default effects play an important role in decisions with a social component, such as allocation decisions. The default effect might be driven by a high "cost" of selecting an additional choice, one that—unlike the outcome choice—can be avoided through passivity. The default choice may also be interpreted as a suggestion on the part of the experimenter or as providing permission to act in a particular way.

The Strategy Method condition shares the same default as the Default NR condition, but almost three-quarters of its participants overcome the default and actively choose to learn the recipients' payoffs. Because participants indicate contingent outcome choices on the same screen, revealing the payoffs serves only to resolve uncertainty ex post. Thus, although the active pursuit of information is consistent with a general desire to resolve uncertainty or with a desire to learn the outcome of one's choice, it is markedly inconsistent with a nonstrategic explanation of the ignorance found in the default conditions. Instead, dictators appear to choose ignorance to avoid having to face an informed decision, but only when they can do so passively.

What is it about facing an informed decision that dictators would want to avoid? Although there is little to suggest that dictators would want to avoid a decision in the AI game, two different motivations might provide a reason to want to avoid a decision in the CI game. Each could explain why participants are willing to choose ignorance passively but are reluctant to do so actively.

First, because their choice determines another person's payoff, dictators may feel obligated to choose the fair outcome, despite having a direct preference for the selfish outcome. Many studies have shown that a sense of responsibility for the welfare of others has an impact on behavior.[10] Dictators may be able to alleviate that responsibility by remaining ignorant, much in the way they can by delegating the decision to an intermediary or by opting out of the allocation decision.[11] Indeed, Krupka and Weber (2013) find that the social consensus among experimental subjects who are given descriptions of the DWK baseline and moral wiggle room game is that remaining ignorant is more socially appropriate than choosing selfishly either in the baseline or while informed of being in the CI game. However, one does wonder how choosing to remain

---

[9] Across all four default conditions, 4 (13%) out of the 30 participants who remained ignorant chose to give up a dollar despite not knowing whether their sacrifice would help the recipient. This behavior, although rare, was found by DWK at a similar rate (14%) and is inconsistent with most models of outcome-based preferences. Although it is compatible with signaling one's selflessness, it is far from clear whether that is the best explanation.

[10] See, for example, Charness (2000) or Charness and Jackson (2009).

[11] See, for example, Coffman (2011), Bartling and Fischbacher (2012), Oexl and Grossman (2013), or Hamman et al. (2010) for the former and Lazear et al. (2012), Dana et al. (2006), or Broberg et al. (2007) for the latter.

ignorant relieves one of responsibility. This would seem to be a form of self-deception, an important topic for future research.

However, whereas unhelpful behavior on the part of the uninformed may be excused, a "sin of commission" may be viewed more harshly than a "sin of omission," and actively pursuing such ignorance may be considered an abrogation of responsibility. This remains to be seen, however, and further research should seek to establish whether or not the active pursuit of ignorance does indeed spoil its responsibility-alleviating properties and is viewed as more inappropriate socially.

A second, related explanation is that dictators are concerned about their image and know that they will be judged differently when ignorant than when informed. However, recipients in this experiment and that of DWK do not observe the dictator's information choice, so remaining ignorant can have no impact on their beliefs. Furthermore, the information and experience of the recipient remained constant across the default manipulations and the second-order beliefs (i.e., beliefs about the recipients' beliefs) reported by the dictators about the proportion of dictators revealing the payoffs varied little with the default conditions. These facts rule out *social* image concern as an explanation for the dictator's chosen ignorance.

Instead, this strategic ignorance has been interpreted as evidence of concern for *self*-image (Bénabou and Tirole 2006, Hamman et al. 2010, Mazar et al. 2008). By not revealing the payoffs, dictators can disregard the fact that they are not helping the recipient while simultaneously acting selfishly. Whereas direct tests of the comparative static predictions of Bayesian models have found little evidence that self-signaling affects giving rates (Grossman 2010), Gneezy et al. (2012) find that some people avoid situations that may harm their self-image. Furthermore, Grossman and van der Weele (2013) show that strategic ignorance is compatible with Bayesian self-signaling and argue, in part relying on the data reported in this paper, that information acquisition choices in social dilemmas provide clear evidence of the impact of self-signaling on unselfish behavior.

The notion that the dictator acts to maintain her self-image is consistent with findings on omission bias, the tendency to favor acts of omission over otherwise equivalent commissions (Ritov and Baron 1990), particularly the finding that harmful omissions are judged less severely than harmful commissions (Spranca et al. 1991). Furthermore, the current findings contribute to a broader literature examining strategic ignorance in various contexts, through which runs the common thread of the analysis of ignorance as a way to avoid to making "wrong" decisions in the future.[12]

---

[12] See, for example, Bénabou and Tirole (2011), Nyborg (2011), Caplin and Leahy (2001), Carillo and Mariotti (2000), Aghion and Tirole (1997), and Crémer (1995).

These results suggest that the critique of the robustness of social preferences implied by the moral wiggle room story may itself not be so robust. Yet I certainly do not wish to suggest that social preferences manifest in all environments. For example, Fershtman et al. (2012) provide striking evidence that introducing a competitive environment makes social preferences evanescent. Matters are far from being settled, and further research will surely follow.

Supplemental Material
Supplemental material to this paper is available at http://dx.doi.org/10.1287/mnsc.2014.1989.

References
Aghion P, Tirole J (1997) Formal and real authority in organizations. *J. Political Econom.* 105(1):1–29.

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

Ariely D, Bracha A, Meier S (2009) Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *Amer. Econom. Rev.* 99(1):544–555.

Bartling B, Fischbacher U (2012) Shifting the blame: On delegation and responsibility. *Rev. Econom. Stud.* 79(1):67–87.

Bénabou R, Tirole J (2006) Incentives and prosocial behavior. *Amer. Econom. Rev.* 96(5):1652–1678.

Bénabou R, Tirole J (2011) Identity, morals and taboos: Beliefs as assets. *Quart. J. Econom.* 126(2):805–855.

Bolton GE, Ockenfels A (2000) ERC: A theory of equity, reciprocity, and competition. *Amer. Econom. Rev.* 90(1):166–193.

Brandts J, Charness G (2011) The strategy versus the direct-response method: A first survey of experimental comparisons. *Experiment. Econom.* 14(3):375–398.

Broberg T, Ellingsen T, Johannesson M (2007) Is generosity involuntary? *Econom. Lett.* 94(1):32–37.

Caplin A, Leahy J (2001) Psychological expected utility theory and anticipatory feelings. *Quart. J. Econom.* 116(1):55–79.

Carillo JD, Mariotti T (2000) Strategic ignorance as a self-disciplining device. *Rev. Econom. Stud.* 67(3):529–544.

Charness G (2000) Responsibility and effort in an experimental labor market. *J. Econom. Behav. Organ.* 42(3):375–384.

Charness G, Dufwenberg M (2006) Promises and partnership. *Econometrica* 74(6):1579–1601.

Charness G, Jackson MO (2009) The role of responsibility in strategic risk-taking. *J. Econom. Behav. Organ.* 69(3):241–247.

Charness G, Rabin M (2002) Understanding social preferences with simple tests. *Quart. J. Econom.* 117(3):817–869.

Coffman LC (2011) Intermediation reduces punishment (and reward). *Amer. Econom. J.: Microeconomics* 3(4):77–106.

Crémer J (1995) Arm's length relationships. *Quart. J. Econom.* 110(2):275–295.

Dana J, Cain DM, Dawes R (2006) What you don't know won't hurt me: Costly (but quiet) exit in a dictator game. *Organ. Behav. Human Decision Processes* 100(2):193–201.

Dana J, Weber RA, Kuang JX (2007) Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Econom. Theory* 33(1):67–80.

Dufwenberg M, Gneezy U (2000) Measuring beliefs in an experimental lost wallet game. *Games Econom. Behav.* 30(2):163–182.

Dufwenberg M, Kirchsteiger G (2004) A theory of sequential reciprocity. *Games Econom. Behav.* 47(2):268–298.

Falk A, Fischbacher U (2006) A theory of reciprocity. *Games Econom. Behav.* 54(2):293–315.

Fehr E, Fischbacher U (2002) Why social preferences matter—The impact of non-selfish motives on competition, cooperation, and incentives. *Econom. J.* 112(478):C1–C33.

Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. *Quart. J. Econom.* 114(3):817–868.

Feiler L (2007) Behavioral biases in information acquisition. Ph.D. thesis, California Institute of Technology, Pasadena.

Fershtman C, Gneezy U, List JA (2012) Equity aversion: Social norms and the desire to be ahead. *Amer. Econom. J.: Microeconomics* 4(4):131–144.

Fischbacher Urs (2007) z-Tree: Zurich toolbox for ready-made economic experiments. *Experiment. Econom.* 10(2):171–178.

Gneezy A, Gneezy U, Riener G, Nelson LD (2012) Pay-what-you-want, identity, and self-signaling in markets. *Proc. Natl. Acad. Sci. USA* 109(19):7236–7240.

Greiner B (2003) An online recruitment system for economic experiments. Kremer K, Macho V, eds. *Forschung und wissenschaftliches Rechnen 2003*, GWDG Bericht 63, (Gesellschaft für wissenschaftliche Datenverarbeitung, Göttingen, Germany), 79–93.

Grossman Z (2010) Self-signaling versus social-signaling in giving. Economics Working Paper 712009, Department of Economics, University of California, Santa Barbara, Santa Barbara. http://ideas.repec.org/p/cdl/ucsbec/712009.html.

Grossman Z, van der Weele J (2013) Self-image and strategic ignorance in moral dilemmas. Economics Working Paper qt0bp6z29t, Department of Economics, University of California, Santa Barbara, Santa Barbara. http://ideas.repec.org/p/cdl/ucsbec/qt0bp6z29t.html.

Hamman JR, Loewenstein G, Weber RA (2010) Self-interest through delegation: An additional rationale for the principal-agent relationship. *Amer. Econom. Rev.* 100(4):1826–1846.

Krupka EL, Weber RA (2013) Identifying social norms using coordination games: Why does dictator game sharing vary? *J. Eur. Econom. Assoc.* 11(3):495–524.

Larson T, Capra CM (2009) Exploiting moral wiggle room: Illusory preference for fairness? A comment. *Judgment Decision Making* 4(6):467–474.

Lazear EP, Malmendier U, Weber RA (2012) Sorting in experiments with application to social preferences. *Amer. Econom. J.: Appl. Econom.* 4(1):136–63.

Madrian B, Shea DF (2001) The power of suggestion: Interia in 401(k) participation and savings behavior. *Quart. J. Econom.* 116(4):1149–1187.

Matthey A, Regner T (2007) Is observed other-regarding behavior always genuine? Jena Economic Research Paper 2007-109, Friedrich-Schiller-University of Jena, Jena, Germany.

Mazar N, Amir O, Ariely D (2008) The dishonest of honest people: A theory of self-concept maintenance. *J. Marketing Res.* 45(6):633–644.

Nyborg K (2011) I don't want to hear about it: Rational ignorance among duty-oriented consumers. *J. Econom. Behav. Organ.* 79(3):263–274.

Oexl R, Grossman ZJ (2013) Shifting the blame to a powerless intermediary. *Experiment. Econom.* 16(3):306–312.

Rabin M (1993) Incorporating fairness into game theory and economics. *Amer. Econom. Rev.* 83(5):1281–1302.

Ritov I, Baron J (1990) Reluctance to vaccinate: Omission bias and ambiguity. *J. Behavioral Decision Making* 3(4):263–277.

Samuelson W, Zeckhauser R (1988) Status quo bias in decision making. *J. Risk Uncertainty* 1(1):7–59.

Spranca M, Minsk E, Baron J (1991) Omission and commission in judgment and choice. *J. Experiment. Soc. Psych.* 27(1):76–105.