

Paying to Be Nice: Consistency and Costly Prosocial Behavior

Ayelet Gneezy

Rady School of Management, University of California, San Diego, La Jolla, California 92093,
agneezy@ucsd.edu

Alex Imas

Department of Economics, University of California, San Diego, La Jolla, California 92093,
aimas@ucsd.edu

Amber Brown

Disney Research Boston, Cambridge, Massachusetts 02142, amber@disneyresearch.com

Leif D. Nelson

Haas School of Business, University of California, Berkeley, Berkeley, California 94720,
leif_nelson@haas.berkeley.edu

Michael I. Norton

Harvard Business School, Harvard University, Boston, Massachusetts 02163, mnorton@hbs.edu

Building on previous research in economics and psychology, we propose that the costliness of initial prosocial behavior positively influences whether that behavior leads to consistent future behaviors. We suggest that *costly* prosocial behaviors serve as a signal of prosocial identity and that people subsequently behave in line with that self-perception. In contrast, *costless* prosocial acts do not signal much about one's prosocial identity, so subsequent behavior is less likely to be consistent and may even show the reductions in prosocial behavior associated with licensing. The results of a laboratory experiment and a large field experiment converge to support our account.

Key words: charitable giving; moral consistency; licensing; field experiment

History: Received July 13, 2010; accepted July 18, 2011, by Brad Barber, Teck Ho, and Terrance Odean, special issue editors. Published online in *Articles in Advance* November 4, 2011.

Introduction

Before he was arrested in June of 2005 for fraud charges, Wall Street financier Alberto Vilar seemed to have it all. Just a few years earlier, Vilar was lunching with the Prince of Wales and having his name engraved in marble at London's Royal Opera House. Given these facts, Vilar's story may seem to be just another tale of an investor falling from grace, except for the fact that his \$200 million fortune was not spent on either sports cars or yachts, but charity—the fraud charges brought against Vilar were based on donation commitments that he did not have the money to fulfill. In fact, Alberto Vilar gave away so much money that just a week before his arrest, he had agreed to pay university tuition for his barber's daughter, after he had already covered her wedding costs.

Although such pathological addiction to charitable giving is uncommon, it remains plausible that a little charitable giving might lead to a little bit more. When does a little prosocial behavior cause someone to see herself as a prosocial type, spurring subsequent simi-

lar behavior? A number of factors likely influence the emergence of such consistency, but we propose one crucial factor: whether the initial prosocial behavior is costly to the individual. We define *costly* prosocial behavior as any action aimed at benefiting others that involves some cost to the agent—which frequently comes in the form of a contribution of time or money to benefit others (Liu and Aaker 2008). We define *costless* prosocial behavior as any behavior that similarly benefits others but does not impose any costs on the agent. Specifically, we propose that when prosocial behavior comes at a cost, it serves as a temporary signal to the self regarding one's prosocial identity and to increased prosocial behavior.

Cost and Consistency

Why would one's behavior serve as a signal of one's identity? Studies in psychology and economics have demonstrated that individuals "consult" with their identities when making choices, preferring to act consistently with the norms prescribed by those identities (Charness et al. 2007, Chen and Li 2009, LeBoeuf

et al. 2010) and avoiding information that may tarnish their self-image (Dana et al. 2007). Similarly, several recent economic models incorporate identity and self-image concerns into decision making, particularly in the context of prosocial behavior. Akerlof and Kranton (2000, 2005) allow for identity to enter as an input into the agent's payoff functions, showing that behavior that clashes with identity decreases the agent's utility, whereas behavior that bolsters identity results in positive payoffs. Bodner and Prelec (2003) propose a multiple-self model where self-image is represented as a probability distribution over possible moral types. Individuals value a secure self-image, where a high probability is placed on possessing qualities that are in line with social norms and personal beliefs. In this framework, agents are unsure of their true moral type and must therefore infer the likelihood of being a particular type from past behavior (see also Bénabou and Tirole 2006, 2011).

Research in psychology has similarly explored cases in which people ascertain their own identity by observing their own behavior (Ariely and Norton 2008, Bem 1972). In a classic study (Bem and McConnell 1970), participants were induced to write essays in favor of issues that were opposed to their previously reported attitudes; subsequently, participants' attitudes regarding those topics were elicited. Importantly, writing essays against their viewpoints led participants to change their attitudes to match their essays. When participants were then asked to recall their *original* attitudes, their reported attitudes were nearly identical to those reported post-manipulation—and significantly different than those recorded before the essays were written. This suggests that participants lacked perfect recall of their previously held attitudes and formed new attitudes through inference based on recent behavior (see also Goethals and Reckman 1973, for a similar result). Indeed, these specific cases are examples of a general tendency to view the self in the present as similar to the self in the past—to see oneself as consistent over time (Greenwald 1980, Ross 1989). In the specific domain of prosocial behavior, these results suggest that people may lack perfect information as to their moral type, such that prosocial behavior may lead them to update their view of themselves: If I behaved prosocially, I must be a prosocial kind of person—someone for whom prosocial behavior provides greater utility—and therefore I will behave more prosocially in the future (see Bénabou and Tirole 2011).

This drive toward consistency between attitudes and behavior, however, is in seeming conflict with other research suggesting that behavior in one direction can license behavior in the opposite direction (Miller and Effron 2010, Monin and Miller 2001), so

that someone who behaves prosocially may feel liberated to subsequently behave antisocially in the future. We propose one critical moderator of when previous behavior leads to either consistency or licensing effects: the costliness of that behavior, which serves as a signal of the informative value of that behavior to one's identity. In short, we suggest that a person engaging in costly moral behavior is likely to infer that she is a moral person: "Why else would I incur a cost to be moral unless I am that kind of person?"

Building on previous literature, we posit that cost induces consistency by making past behavior more diagnostic about oneself. In the context of economic models of prosocial behavior, where the act serves as a signal of one's moral type, the logic of Bénabou and Tirole (2004) suggests that the lower the cost of the signal, the less informative it is regarding one's moral type. Indeed, in two classic consistency paradigms—the foot in the door paradigm and the original cognitive dissonance paradigm—the cost of engaging in some moral behavior is crucial in producing subsequent consistency, though these theorists have not explicitly identified cost as a key mechanism in their findings.

Freedman and Fraser's (1966) foot in the door paradigm—in which inducing someone to agree to perform a small prosocial act is used to facilitate agreement to a larger one—had the goal of getting homeowners to agree to allow five or six men into their homes for two hours to classify all of their household products by going through their cupboards and closets. In one version, the caller merely asked if the homeowners were willing to answer eight questions about consumer goods; in the other, the homeowners actually had to answer these questions. In our view, the first condition is relatively costless—agreeing to do the caller a favor and answer questions in the future—whereas the second is more costly—actually taking the time to answer those questions. Our account would suggest that only in the latter case would homeowners be more likely to help further by assenting to the larger request. The results of Freedman and Fraser (1966) are consistent with our proposition: Whereas just 33.3% complied in the former condition, some 52.8% did in the latter—when the initial prosocial behavior was costly. In addition, a meta-analysis of studies examining the foot in the door paradigm showed a significant positive relationship between the level of cost—or involvement—required of the participant by the initial task and the rate of compliance in agreeing to a second task (Burger 1999). Similarly, in Festinger and Carlsmith's (1959) dissonance paradigm, an actor who behaved immorally by telling an unsuspecting person that a boring task (turning wooden knobs) was quite fun subsequently professed enjoyment of

that task to bring preferences in line with otherwise antisocial behavior. Later theorists explicitly defined such aversive consequences to the self—the cost of one’s poor behavior—as a critical factor for producing the kinds of consistency observed in dissonance paradigms (Cooper and Fazio 1984, Steele 1988, Stone and Cooper 2001, Thibodeau and Aronson 1992).

Importantly for our account, many of the paradigms used to demonstrate licensing effects have utilized initially costless tasks—often via hypothetical scenarios—and then observed changes in subsequent behavior. For instance, people asked to imagine engaging in community service were subsequently more likely to treat themselves to frivolous products (Khan and Dhar 2006), people first asked to allocate hypothetical money to a charity subsequently allocated less real money to charity than did those in a control condition (Falk and Zimmermann 2010), people given an opportunity to express their endorsement of Barack Obama subsequently allocated more money to an organization serving Whites (Effron et al. 2009), people who wrote positive stories about themselves subsequently pledged less money to charity (Sachdeva et al. 2009), and people who purchased green products using house money were less altruistic in a subsequent dictator game (Mazar and Zhong 2010). Because these paradigms involve hypothetical or costless initial decisions, our account would not predict the emergence of consistency; instead, we propose that the presence of cost in the initial behavior is critical for consistency to emerge.

How do behavior, costliness, and identity work in concert to produce either licensing or consistency? Consider the following highly stylized example of an individual with a medium level of prosocial identity, who engages in either costless or costly prosocial behavior. If she is a “5” on a 10-point scale of prosocial identity and behaves like a “7” person, the classic licensing prediction would be that she is likely to subsequently behave in a “3” fashion to net an overall “5.” In our view, however, that prediction is contingent on the initial behavior being costless and therefore not influencing her prosocial identity. In contrast, if she behaves like a “7” person and that behavior is costly, she may update her view of herself (to, in this stylized example, a “6” person) and match her subsequent behavior to this new more prosocial identity—leading to consistency. Taken together, our account makes two predictions: Costless behavior should *not* change one’s prosocial identity and therefore should lead to licensing effects, whereas costly behavior should both change one’s prosocial identity and lead to subsequent behavior consistent with that prosocial identity.

We present the results of a laboratory experiment and a field experiment that test these predictions. We

designed Experiment 1 to test whether costly prosocial behavior is more likely to lead to subsequent moral behavior than similar costless behavior and to see if this pattern of behavior is driven by a positive shift in individuals’ perceptions of their prosociality. In Experiment 2 we move to a field setting, examining how purchasing products with a prosocial component impacts subsequent other-regarding behavior; because purchasing products is by definition a costly behavior, we expected to observe consistency rather than licensing.

Experiment 1—Costly and Costless Prosocial Behavior in the Lab

We sought to test the idea that costly forms of altruism would lead to subsequent increases in prosocial behavior, whereas less costly forms of altruism would lead to licensing effects and subsequent decreases in prosocial behavior. In addition, as outlined above, we expected costly prosocial behavior to send a signal to the self regarding one’s altruistic preferences, leading to an increase in self-reported prosocial identity. In contrast, because costless altruism does not send a strong signal, we expected no change in prosocial identity.

To manipulate the cost of altruism, some participants had part of their payment donated to charity (costly altruism); other participants learned that a similar donation had been given on their behalf, but that it was not deducted from their payment (costless altruism).

To measure subsequent prosocial behavior we created a task in which people could either lie to benefit themselves or tell the truth to benefit an unspecified other. In addition, we measured participants’ perceived prosocial identity to examine the effect that one’s initial prosocial behavior has on her self-perceptions of prosociality. We predicted that those engaging in costly altruism would rank higher on this measure, relative to participants in the control and costless prosocial treatments.

Participants played a two-player game and were assigned the role of either Sender or Receiver (Gneezy 2005). In this game the Sender has private information. The Receiver needs to make a choice based on a message conveyed by the Sender. Payoffs for both players depend on a choice made by the Receiver. We designed the game such that the Sender had a monetary incentive to deceive the Receiver by sending a false message. To measure the effect of costly and costless altruism on such deceptive behavior, Senders made a mandatory donation to the Make-a-Wish Foundation that was either deducted from their payment (costly treatment) or was made independent of their payment (costless treatment). Receivers did

Table 1 Payoff Scheme by Treatment

| Treatment | Option | Sender \$ | Receiver \$ | Thank you \$ | Donation \$ | Sender net \$ | <i>N</i> |
|-------------------|--------|-----------|-------------|--------------|-------------|---------------|----------|
| Control | X | 30 | 25 | 5 | — | 35 | 62 |
| | Y | 25 | 30 | 5 | — | 30 | |
| Costly Donation | X | 30 | 25 | 3 | 2 | 33 | 49 |
| | Y | 25 | 30 | 3 | 2 | 28 | |
| Costless Donation | X | 30 | 25 | 5 | 2 | 35 | 57 |
| | Y | 25 | 30 | 5 | 2 | 30 | |

Notes. Summary of treatment parameters (Experiment 1). The “Option” column refers to the Option, X or Y, that determines the payments to the Sender and Receiver. The “Sender \$” and “Receiver \$” columns denote the total amount each player would gain, respectively, depending on the option used for payment, not including the donation or “thank you” payment. The “thank you \$” column refers to the payment enclosed in the Senders’ envelopes at the start of the experiment. The “Donation \$” column lists the donation made on behalf of the Senders. The “Sender net \$” column denotes the net amount the Sender would gain after accounting for the donation and “thank you” payment, depending on the option used for payment. The column “*N*” denotes the number of participants per treatment.

not have this option and were not informed of its availability to Senders.

Methods

Undergraduates ($N = 168$) participated in this experiment as part of a course in intermediate economics. The sample included all students attending class the day of the experiment. Neither the instructor of the course nor students taking it had any information regarding the objective of the experiment. In addition, students were informed that their decisions and responses would be kept confidential and analyzed at the aggregate level.

Those assigned to the role of Sender were part of a single class, and those assigned the role of Receiver were part of another. Senders were then randomly assigned to one of three treatments: Costless Donation, Costly Donation, and Control. In all three treatments, participants were first given an envelope containing a brief introduction and a note indicating their \$5 “thank you” payment for participating in the experiment. In the Control treatment, five \$1 bills were enclosed in the envelope. Participants in the Costly Donation treatment were informed that \$2 had been automatically deducted from their payment and would be donated to charity—three \$1 bills were enclosed in their envelopes. Participants in the Costless Donation treatment were told that \$2 would be donated on their behalf *independently* of their payment and that this donation would not affect their final earnings; as in the Control treatment, five \$1 bills were enclosed in their envelopes.¹ The charity in both

donation treatments was the Make-a-Wish Foundation. At the end of the experiment, we made the donations on the behalf of Senders via the Make-a-Wish Foundation website.

Once participants had familiarized themselves with the first set of instructions, a second envelope was passed out. To assess whether our costliness manipulation impacted participants’ self-perceptions, as our account predicts, Senders in all three treatments then completed two items on 5-point scales (1: *not at all* to 5: *very much*) assessing the extent to which they saw themselves as “helpful” and “selfish”—measuring their perceived prosocial identity (Aquino and Reed 2002, Batson et al. 1978, Grant et al. 2008). The two items were significantly correlated, $r(168) = -0.32$, $p < 0.001$; we therefore reverse-coded “selfish” and averaged the two items to create a composite measure of prosocial identity.

Next, Senders were given the Sender–Receiver game instructions in which we informed them that 1 in 10 participants would be randomly chosen to be matched with a Receiver in another class and paid according to the instructions.² Participants assigned to the role of Receivers did not receive a first envelope with the donation manipulation nor the prosocial identity measures; they were simply told that they would be paid based on their decisions. Both sets of instructions (see the online appendix, available at <http://rady.ucsd.edu/faculty/directory/ayelet-gneezy/>) specified that there were two possible payment outcomes to the experiment: Option X and Option Y. Senders had full information about the monetary payoffs for each option, whereas Receivers

¹ This forced-compliance design—where every participant in a given treatment is either obligated to donate or not—was used to avoid issues pertaining to self-selection and is analogous to those typically employed in investigations of altruism (Andreoni 1990, Dunn et al. 2008). Neuroeconomic studies demonstrate that reward regions are activated when individuals donate to charity whether by choice or not (e.g., Harbaugh et al. 2007, Moll et al. 2006).

² The practice of randomly selecting participants for payment has been used in several investigations in both psychology and economics (e.g., Erat and Gneezy 2012, Mazar and Zhong 2010) and is conceptually similar to alternate schemes that have been used successfully (Laury 2006). The intent of this practice is to induce participants to treat *every* decision as consequential, precisely because *any* decision might be so—thereby creating an incentive-compatible environment.

had no information about the payoffs.³ We informed Senders that if the Receiver chose a number that corresponded to the last digit of the Sender’s personal identification number, both players would be paid according to payment Option Y. If the Receiver chose any other number, both players would be paid according to Option X. Importantly, Option Y meant that the Receiver gets more money than the Sender, whereas Option X resulted in the opposite payoff scheme, such that the Sender gets more money than the Receiver (see Table 1 for a complete payoff scheme).

Senders sent a single message to the Receiver phrased as: “Choosing ___ will earn you more money than any other number.” The Sender chose which number to write in the blank. If the Sender indicated the correct number then the message was honest, and if followed, would benefit the Receiver to the detriment of the Sender. If the Sender chose any incorrect number then the message was dishonest, and the Sender would benefit to the detriment of the Receiver.⁴ Given the payoff scheme, the Senders had a monetary incentive to deceive their counterparts.

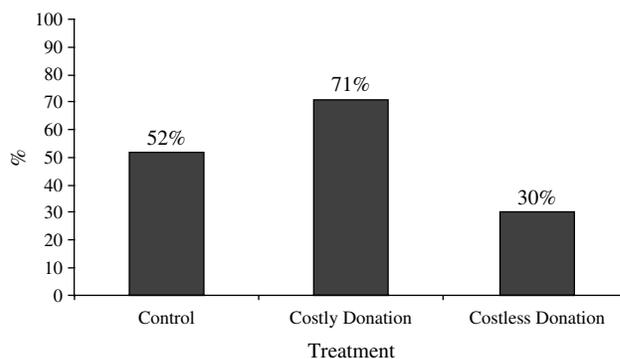
We informed all participants that neither the Sender nor Receiver would ever know the identity of the player with whom they were matched.

Results and Discussion

Prosocial Identity. The three treatments differed in the extent to which people reported feeling prosocial, $F(2, 167) = 8.90, p < 0.001$. Specifically, we observed the predicted effect of cost on self-perceptions: Participants in the Costly treatment reported feeling more prosocial ($M = 3.85, SD = 0.86$) than those in the Control treatment ($M = 3.27, SD = 0.82$), $z = 3.60, p < 0.001$, and those in the Costless treatment ($M = 3.26, SD = 0.75$), $z = 3.60, p < 0.001$.⁵ Participants in the Control and Costless treatments did not differ in their reports of prosocial identity, $z = 0.37, p = 0.71$.

Truth Telling. Our model suggests that the differences observed in participants’ prosocial identity should be accompanied by shifts in prosocial behavior. A comparison of the percentage of participants who told the truth in each treatment revealed a significant difference, Pearson’s $\chi^2(2) = 18.35, p < 0.001$ (Figure 1). More importantly and consistent with our predictions, participants in the Costly Donation treatment were substantially more likely to send a truthful message (71%) than participants in the Control

Figure 1 Percentage of Participants Who Told the Truth in Experiment 1



treatment (52%), $z = 2.11, p = 0.04$, and the Costless treatment (30%), $z = 4.25, p < 0.001$. In addition, participants in the Control treatments were more likely to tell the truth than those in the Costless treatment, $z = -2.40, p = 0.02$.⁶

As predicted, participants in the Costly treatment ranked significantly higher on the prosocial identity measure than Control and Costless participants. Costly participants were more likely to subsequently engage in prosocial behavior, relative to Control and Costless participants. Participants in the Costless condition did not experience an increase in their prosocial identity despite having been associated with the same charitable organization and donation size. In addition, Costless participants’ behavior in the Sender–Receiver game was consistent with moral licensing—they were significantly more likely to deceive their counterparts than Control and Costly participants.

Mediation Analysis. We next examine whether the effect of cost on truth-telling behavior is mediated by a shift in prosocial identity in the Costly and Costless Donation treatments, using the technique proposed by Baron and Kenny (1986) in conjunction with the bootstrapping method of Shrout and Bolger (2002). Significance coefficients were calculated using bias-corrected confidence intervals, as these have been shown to provide more accurate estimates for small samples (Efron and Tibshirani 1993). First, costliness predicted both prosocial identity ($\beta = 0.35, p < 0.001$) and truth telling ($\beta = 0.42, p < 0.001$). The effect of

³ This procedure helps avoid the type of message deterioration observed in Blume et al. 2001 (see also Brandts and Charness 2003).

⁴ To minimize the potential for sophisticated deception, we used a richer message space composed of 10 possible messages in order (Sutter 2009).

⁵ The z -values and corresponding p -values were calculated using the nonparametric Mann-Whitney test.

⁶ This experiment was conducted as a follow-up to a previous experiment that used a very similar paradigm but with larger incentives to lie. Given these incentives, we found that in general participants lied more often, but that Costly participants were substantially more likely to send a truthful message (41%) than Costless participants (18%) and Control participants (21%) ($z = 2.60, p = 0.01$; $z = 2.26, p = 0.02$, respectively). Costless and Control participants were equally truthful, $z = -0.38, p = 0.70$. In short, because lying was so common, the ceiling effect allowed us to detect the presence of consistency but not licensing. As a result, we adjusted the incentives before conducting Experiment 1.

costliness on truth telling was reduced (to $\beta = 0.36$, $p < 0.001$) when prosocial identity was included in the equation, and prosocial identity was a significant predictor of truth telling ($\beta = 0.17$, $p = 0.02$). The 95% bias-corrected confidence intervals for the size of the indirect effect excluded zero (0.015, 0.114), suggesting a significant indirect effect; as predicted, prosocial identity mediated the effect of cost on truth telling.

Experiment 2—Costly Prosocial Behavior in the Field

Having established the positive impact of an initial costly prosocial behavior on individuals' self-perceptions of prosocial identity and subsequent prosocial behavior, we next ran a large field experiment to test our predictions. We conducted the experiment at a ride in a large American amusement park. Visitors were photographed while on the ride and were then given the opportunity to purchase their photo. To test the effect of prosocial behavior on subsequent prosocial choices, we randomly assigned participants to one of two treatments that differed with respect to whether the photo was sold with a charitable-giving promotion. We interpret purchasing a photo with a charitable-giving promotion as costly prosocial behavior because individuals incur a monetary cost to benefit a social cause. We then assessed whether subsequent purchases made at that store were more likely to be for oneself (a selfish behavior) or gifts for others (a prosocial behavior; see Dunn et al. 2008). This field setting allows us to test for the emergence of consistency or licensing in a naturalistic setting: Having behaved prosocially by buying a product with a charitable component, will people feel licensed to then buy subsequent products for themselves, or, as our account holds, will the act of engaging in costly prosocial behavior spur them to behave prosocially in their subsequent purchases?

Methods

The study was conducted at a popular ride within a large U.S. amusement park over the course of nine days. At the exit of the ride, visitors had the option to purchase a 5×7 printed photograph taken during their ride.

To test our proposition that an initial costly prosocial behavior will increase the likelihood that one would subsequently behave prosocially, we manipulated whether the photo was sold under a charitable-giving promotion or under a similar promotion without a charitable component. On some days the photo was sold such that half of the price was donated to a charity (a major patient-support organization); on others, photos were sold without the charitable element. Participants chose whether to buy

the photo and then walked through a retail area that contained other merchandise. Surveyors intercepted a random sample of photo purchasers at the exit of this retail area and asked them three yes/no questions: "Did you buy a photo?," "Did you buy any other merchandise?," and in reference to the latter, "Are any of those purchases going to be gifts for others?" Responses to the final question served as our primary dependent measure; buying a gift for someone is a prototypical prosocial behavior (Dunn et al. 2008).⁷

Our analysis is based on respondents who bought merchandise in the retail area ($N = 363$) and who fall into one of four groups depending on whether they purchased a photo and whether the photo was sold under a charitable or a standard promotion. The experiment was conducted over nine days; each day was randomly assigned to one of the two treatments. Responses were collected by nine surveyors.

Results and Discussion

We analyzed the likelihood that a customer subsequently bought merchandise for someone else (versus for themselves) as a function of whether they were exposed to the charity promotion and whether they bought the initial photo (see Table 2). Our binary dependent variable was *Gift Purchases for Others*. Our independent variables, all binary, were *Photo Purchase* (yes or no), *Prosocial Promotion* (yes or no), and the interaction of the two. In addition, we controlled for surveyor effects by including dummy variables for each individual surveyor. Because we have a binary dependent variable and multiple binary independent variables, we ran a logit regression for our analysis. In addition, we ran two more regressions to further explore the effect of *Photo Purchase* on *Gift Purchases for Others*. Specifically, we separately regressed *Gift Purchases for Others* on *Photo Purchase* conditional on (a) the *Prosocial Promotion* being present (*Prosocial Promotion* = 1), as well as (b) not being present (*Prosocial Promotion* = 0). Dummy variables for individual surveyors were included in both regressions. Results are reported in Table 3.

Neither photo purchasing ($\beta = -0.18$, $p = 0.67$) nor the prosocial promotion ($\beta = -0.20$, $p = 0.62$) influenced subsequent gift buying. Importantly, however, we observed a marginally significant interaction ($\beta = 1.04$, $p = 0.05$). When participants were exposed to the prosocial cue (*Prosocial Promotion* = 1), those

⁷The survey also contained questions asking people how much they enjoyed the ride, their likelihood of recommending a vacation to the amusement park, how participants found out about the prices, how many people were in their traveling party, whether this was their first visit to the amusement park, their home zip code, their age, and their gender. None of these variables was determined to be relevant to the analysis, so they were not included.

Table 2 Percentage of Customers Purchasing Merchandise as a Gift for Someone Else

| | No charity (<i>N</i> = 167) | Charity (<i>N</i> = 196) |
|---------------------------------------|---------------------------------|------------------------------|
| Photo Purchaser (<i>N</i> = 185) | 41% | 56% |
| Non Photo Purchaser (<i>N</i> = 178) | 45% | 40% |

Note. Average percentages of customers purchasing gifts for others in Experiment 2, adjusted for surveyor effects.

who purchased a photo were *more* likely than non-purchasers to buy subsequent merchandise for others ($\beta = 0.82, p = 0.04$), consistent with our costly altruism account. Although all participants were exposed to the prosocial cue, only those who engaged in a costly prosocial behavior demonstrated increases in subsequent prosocial behavior. When participants were *not* exposed to the prosocial cue (*Prosocial Promotion* = 0), photo purchasers and nonpurchasers were equally likely to buy merchandise for others ($\beta = -0.45, p = 0.43$).

It is possible, however, that rather than the costly initial purchase of a charity photo causing subsequent prosocial purchases, charity photo buyers were simply more dispositionally altruistic and therefore more likely to purchase subsequent merchandise for others. When we asked participants about their *initial* photo purchase, those in the charity treatment were no more likely to say it was for someone else than were people in the noncharity treatment (32% versus 30%, $\beta = 0.03, p = 0.73$). It appears as though our observations of consistent prosocial behavior are not due to a pre-existing inclination for gift-buying but rather are a direct result from the costly act of purchasing.

Table 3 Results of Logit Regression

| | β^a | β^b | β^c |
|---|------------------|------------------|-----------------|
| Photo Purchase | -0.18 (0.44) | | |
| Prosocial Promotion | -0.20 (0.41) | | |
| Photo Purchase \times Prosocial Promotion | 1.04** (0.53) | | |
| Photo Purchase: Prosocial Promotion = 1 | | 0.82** (0.41) | |
| Photo Purchase: Prosocial Promotion = 0 | | | -0.45 (0.56) |

Note. Results from Experiment 2, beta coefficients from three logit regressions with standard errors are in parentheses.

^aLogit regression of *Gift Purchases for Others* on *Photo Purchase*, *Prosocial Promotion*, *Photo Purchase* \times *Prosocial Promotion* and surveyor dummy variables (*N* = 363).

^bLogit regression of *Gift Purchases for Others* on *Photo Purchase* conditional on *Prosocial Promotion* = 1 and surveyor dummy variables (*N* = 196).

^cLogit regression of *Gift Purchases for Others* on *Photo Purchase* conditional on *Prosocial Promotion* = 0 and surveyor dummy variables (*N* = 167).

**Denotes significance of 5%.

Because of the field nature of this experiment, we were unable to control for a number of additional factors that might have influenced our results. One potential factor worthy of consideration is the public nature of purchase decisions. In Experiment 1 participants made their decisions to behave prosocially (or not) in relative privacy, but the nature of an amusement park suggests that many of our respondents completed both their initial photo purchases and subsequent merchandise purchases in the presence of others—in particular, their spouses and children. Our account is agnostic as to whether the presence of others would make people more or less prosocial in general, but we could imagine that having behaved prosocially in public makes subsequent public prosocial behavior even more likely. Although we have focused on the impact of incurring economic costs on prosocial identity, the *social* costs of appearing inconsistent in the eyes of others may serve as another source of “cost” that leads to consistency.

General Discussion

When do initial prosocial acts increase the likelihood of subsequent prosocial acts? Drawing on self-perception theory (e.g., Bem 1972, Goethals and Reckman 1973), we proposed that when recent prosocial behavior was personally costly, people would interpret that behavior as a signal of their prosocial identity; we predicted and demonstrated that having drawn that inference, they would be more likely to subsequently behave prosocially. Prosocial behavior that comes at lower cost, in contrast, offers a more ambiguous signal: prosocial behavior is clearly positive, yet because it came at no cost, it is less likely to be judged as diagnostic of one’s prosocial disposition. Under these circumstances the positive act does not affect individuals’ self-perceptions, resulting in a reduction in subsequent prosocial behavior (“...so I am going to be a bit more greedy now.”).

Two experiments support this reasoning. In the first experiment, we demonstrate that participants who made a costly donation felt more prosocial and behaved more honestly than did participants in control conditions. Furthermore, as predicted, those who made a costless donation did not update their prosocial identity and behaved less honestly than participants in the Control treatment—a licensing effect. We followed up with a field experiment in which visitors to a large theme park could purchase a photo of themselves. Visitors who purchased photos on days in which half of the price was donated to charity were more likely to purchase subsequent items in the store as gifts for others than were visitors on days without this charitable component, again suggesting that costly prosocial behavior is likely to lead to further prosocial behavior.

These findings have implications for understanding when and why consistency and licensing may emerge in the real world. In our laboratory experiment, we are able to demonstrate the isolated effect of causing a prosocial outcome without the incurrance of a personal cost. We suggest that this carefully orchestrated independence of cost from consequence makes understanding how the phenomenon will operate in a more naturalistic environment more complicated. Indeed, initial prosocial behavior in the real world often does come with an associated cost—we often have to put our money and time where our mouth is when behaving prosocially, from paying for prosocial products, as in our field study, to donating our money and volunteering our time to charities. Both earlier research on licensing as well as the studies we report here occur within very brief windows of time—even the shoppers in our field study move fairly quickly from their first purchase to their second, and in lab studies the time between acts is even shorter. As a result, more research is needed to examine the time course of licensing and consistency, examining when and why behaving prosocially leads people to behave either poorly or well, not just over the course of an hour but over the course of their lives.

We note that our instantiation of cost is always both binary and monetary: participants either pay or not. Treating cost as a dichotomous variable is useful for our experiments, but cost clearly exists on a continuum: actions are not either costless or costly, but rather vary in their costliness. Our account suggests that as actions become relatively more costly, consistency is more likely to emerge. In addition, although we have operationalized “cost” in purely monetary terms in the present investigation, costs can come in many forms—greater investment of time, for instance, likely also serves as a costly signal to the self. More abstractly, we would speculate that public statements in support of some cause would be experienced as more costly than similar private statements, because of the potential reputational costs of failing to behave in line with one’s public pronouncements. In other words, our account suggests that any actions that are perceived as costly by the actor—whether or not those costs involve money—are more likely to produce consistency than licensing. Again, future research is needed to explore how relative costs, and the form those costs take, lead to either licensing or consistency.

Finally, the current work also informs a decades-long theoretical dispute across both psychology and economics over whether social preferences—selfless behavior toward another—truly exist (Batson et al. 1997, Cialdini et al. 1997, Levitt and List 2007). It appears that, in part, this debate is rooted in the

different orientations toward the study of prosocial behavior, with psychologists generally exploring why people do not help enough and economists exploring why, given a lack of incentives, people help at all (Ariely and Norton 2007, Charness and Haruvy 2002, Charness and Rabin 2002, Darley and Batson 1973, Fehr and Schmidt 1999, Loewenstein et al. 1989). Consequently, each discipline developed distinct approaches to study social preferences. For instance, psychologists have viewed the emergence of consistent prosocial behavior primarily as a function of internal moral standards and processes such as self-perception, where the self-concept is partially inferred from past behavior, resulting in moral consistency over time (Bem 1972, Festinger 1957, Heider 1958); economists, in contrast, have sought to view the emergence and perpetuation of prosocial behavior as driven by utility maximization, offering data and models for how seemingly suboptimal selfless behavior can in fact be explained as utility maximizing (Andreoni 1990, Bénabou and Tirole 2011). By combining both economic and psychological approaches to the emergence of consistent prosocial behavior, we show that the two approaches in fact converge: When behavior is costly to the self—as in classic psychology studies such as Freedman and Fraser (1966)—that cost serves to make the behavior an informative input in the inference of one’s moral disposition, and that changed self-perception also changes one’s expected utility for future similar behaviors, leading to consistency.

Acknowledgments

The authors thank three anonymous reviewers for their helpful comments. The authors also thank Terrance Odean (special issue coeditor) for his invaluable guidance. This work was partially supported by the National Science Foundation through a Graduate Research Fellowship to Alex Imas.

References

- Andreoni, J. 1990. Impure altruism and donations to public goods: A theory of warm-glow giving. *Econom. J.* **100**(401) 464–477.
- Akerlof, G. A., R. E. Kranton. 2000. Economics and identity. *Quart. J. Econom.* **115**(3) 715–753.
- Akerlof, G. A., R. E. Kranton. 2005. Identity and the economics of organizations. *J. Econom. Perspect.* **19**(1) 9–32.
- Aquino, K., A. Reed II. 2002. The self-importance of moral identity. *J. Personality Soc. Psych.* **83**(6) 1423–1440.
- Ariely, D., M. I. Norton. 2007. Psychology and experimental economics: A gap in abstraction. *Current Directions Psych. Sci.* **16**(6) 336–339.
- Ariely, D., M. I. Norton. 2008. How actions create—not just reveal—preferences. *Trends Cognitive Sci.* **12**(1) 13–16.
- Baron, R. M., D. A. Kenny. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *J. Personality Soc. Psych.* **51**(6) 1173–1182.

- Batson, C. D., J. S. Coke, M. L. Jasnosi, M. Hanson. 1978. Buying kindness: Effects of an extrinsic incentive for helping on perceived altruism. *Personality Soc. Psych. Bull.* 4(1) 86–91.
- Batson, C., K. Sager, E. Garst, M. Kang, K. Rubchinsky, K. Dawson. 1997. Is empathy-induced helping due to self-other merging? *J. Personality Soc. Psych.* 73(3) 495–509.
- Bem, D. J. 1972. Self-perception theory. L. Berkowitz, ed. *Advances in Experimental Social Psychology*, Vol. 6. Academic Press, New York, 1–62.
- Bem, D. J., H. K. McConnell. 1970. Testing the self-perception explanation of dissonance phenomena: On the salience of premanipulated attitudes. *J. Personality Soc. Psych.* 14(1) 23–31.
- Bénabou, R., J. Tirole. 2004. Willpower and personal rules. *J. Political Econom.* 112(4) 848–886.
- Bénabou, R., J. Tirole. 2006. Incentives and prosocial behavior. *Amer. Econom. Rev.* 96(5) 1652–1678.
- Bénabou, R., J. Tirole. 2011. Identity, morals, and taboos: Beliefs as assets. *Quart. J. Econom.* 126(2) 805–855.
- Blume, A., D. V. DeJong, Y. G. Kim, G. B. Sprinkle. 2001. Evolution of communication with partial common interest. *Games Econom. Behav.* 37(1) 79–120.
- Bodner, R., D. Prelec. 2003. Self-signaling and diagnostic utility in everyday decision making. I. Brocas, J. D. Carillo, eds. *The Psychology of Economic Decisions, Vol. 1: Rationality and Well-Being*. Oxford University Press, Oxford, UK, 105–126.
- Brandts, J., G. Charness. 2003. Truth or consequences: An experiment. *Management Sci.* 49(1) 116–130.
- Burger, J. 1999. The foot-in-the-door compliance procedure: A multiple-process analysis and review. *Personal Soc. Psych. Rev.* 3(4) 303–332.
- Charness, G. B., E. Haruvy. 2002. Altruism, fairness, and reciprocity in a gift-exchange experiment: An encompassing approach. *Games Econom. Behav.* 40(2) 203–231.
- Charness, G. B., M. Rabin. 2002. Understanding social preferences with simple tests. *Quart. J. Econom.* 117(3) 817–869.
- Charness, G. B., L. Rigotti, A. Rustichini. 2007. Individual behavior and group membership. *Amer. Econom. Rev.* 97(4) 1340–1352.
- Chen, Y., X. Li. 2009. Group identity and social preference. *Amer. Econom. Rev.* 99(1) 431–457.
- Cialdini, R., S. Brown, B. Lewis, C. Luce, S. Neuberg. 1997. Reinterpreting the empathy-altruism relationship: When one into one equals oneness. *J. Personality Soc. Psych.* 73(3) 481–494.
- Cooper, J., R. H. Fazio. 1984. A new look at dissonance theory. L. Berkowitz, ed. *Advances in Experimental Social Psychology*, Vol. 17. Academic Press, New York, 229–266.
- Dana, J., R. A. Weber, J. X. Kuang. 2007. Exploiting moral wriggle room: Experiments demonstrating an illusory preference for fairness. *Econom. Theory* 33(1) 67–80.
- Darley, J. M., D. Batson. 1973. From Jerusalem to Jericho: A study of situational and dispositional variables in helping behavior. *J. Personality Soc. Psych.* 27(1) 100–108.
- Dunn, E. W., L. B. Aknin, M. I. Norton. 2008. Spending money on others promotes happiness. *Science* 319(21) 1687–1688.
- Efron, B., R. J. Tibshirani. 1993. *An Introduction to the Bootstrap*. Chapman & Hall/CRC, New York.
- Efron, D., J. S. Cameron, B. Monin. 2009. Endorsing Obama licenses favoring Whites. *J. Experiment. Soc. Psych.* 45(3) 590–593.
- Erat, S., U. Gneezy. 2012. White lies. *Management Sci.* Forthcoming.
- Falk, A., F. Zimmermann. 2010. Preferences for consistency. Mimeo, University of Bonn, Bonn, Germany.
- Fehr, E., K. M. Schmidt. 1999. A theory of fairness, competition and co-operation. *Quart. J. Econom.* 114(3) 817–868.
- Festinger, L. 1957. *A Theory of Cognitive Dissonance*. Stanford University Press, Stanford, CA.
- Festinger, L., J. M. Carlsmith. 1959. Cognitive consequences of forced compliance. *J. Abnormal Soc. Psych.* 58(2) 203–210.
- Freedman, J. L., S. C. Fraser. 1966. Compliance without pressure: The foot-in-the-door technique. *J. Personality Soc. Psych.* 4(2) 196–202.
- Gneezy, U. 2005. Deception: The role of consequences. *Amer. Econom. Rev.* 95(1) 384–395.
- Goethals, G. R., R. F. Reckman. 1973. The perception of consistency in attitudes. *J. Experiment. Soc. Psych.* 9(6) 491–501.
- Grant, A. M., J. E. Dutton, B. Rosso. 2008. Giving commitment: Employee support programs and the prosocial sensemaking process. *Acad. Management J.* 51(5) 898–918.
- Greenwald, A. G. 1980. The totalitarian ego: Fabrication and revision of personal history. *Amer. Psychologist* 35(7) 603–618.
- Harbaugh, W. T., U. Mayr, D. R. Burghart. 2007. Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* 316(5831) 1622–1625.
- Heider, F. 1958. *The Psychology of Interpersonal Relations*. John Wiley & Sons, New York.
- Khan, U., R. Dhar. 2006. Licensing effect in consumer choice. *J. Marketing Res.* 43(2) 259–266.
- Laury, S. K. 2006. Pay one or pay all: Random selection of one choice for payment. Working paper, Georgia State University, Atlanta.
- LeBoeuf, R. A., E. Shafir, J. B. Bayuk. 2010. The conflicting choices of alternating selves. *Organ. Behav. Human Decision Processes* 111(1) 48–61.
- Levitt, S. D., J. A. List. 2007. What do laboratory experiments measuring social preferences reveal about the real world? *J. Econom. Perspect.* 21(2) 153–174.
- Liu, W., J. Aaker. 2008. The happiness of giving: The time-ask effect. *J. Consumer Res.* 35(3) 543–557.
- Loewenstein, G., L. Thompson, M. Bazerman. 1989. Social utility and decision making in interpersonal contexts. *J. Personality Soc. Psych.* 57(3) 426–441.
- Mazar, N., C. Zhong. 2010. Do green products make us better people? *Psych. Sci.* 21(4) 494–498.
- Miller, D. T., D. A. Efron. 2010. Psychological license: When it is needed and how it functions. M. P. Zanna, J. M. Olson, eds. *Advances in Experimental Social Psychology*, Vol. 43. Elsevier, San Diego, 115–155.
- Moll, J., F. Krueger, R. Zahn, M. Pardini, R. Oliveira-Souza, J. Grafman. 2006. Human fronto-mesolimbic networks guide decisions about charitable donation. *Proc. Natl. Acad. Sci. USA* 103(42) 15623–15628.
- Monin, B., D. T. Miller. 2001. Moral credentials and the expression of prejudice. *J. Personality Soc. Psych.* 81(1) 33–43.
- Ross, M. 1989. The relation of implicit theories to the construction of personal histories. *Psych. Rev.* 96(2) 341–357.
- Sachdeva, S., R. Iliev, D. L. Medin. 2009. Sinning saints and saintly sinners: The paradox of moral self-regulation. *Psych. Sci.* 20(4) 523–528.
- Shrout, P. E., N. Bolger. 2002. Mediation in experimental and non-experimental studies: New procedures and recommendations. *Psych. Methods* 7(4) 422–445.
- Steele, C. M. 1988. The psychology of self-affirmation: Sustaining the integrity of the self. L. Berkowitz, ed. *Advances in Experimental Social Psychology*, Vol. 21. Academic Press, San Diego, 261–302.
- Stone, J., J. Cooper. 2001. A self-standards model of cognitive dissonance. *J. Experiment. Soc. Psych.* 37(3) 228–243.
- Sutter, M. 2009. Deception through telling the truth? Experimental evidence from individuals and teams. *Econom. J.* 119(534) 47–60.
- Thibodeau, R., E. Aronson. 1992. Taking a closer look: Reasserting the role of the self-concept in dissonance theory. *Personality Soc. Psych. Bull.* 18(5) 591–602.