Journal of Experimental Social Psychology 70 (2017) 198–203

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Case Report

Moral traps: When self-serving attributions backfire in prosocial behavior

Stephanie C. Lin *, Julian J. Zlatev, Dale T. Miller

Stanford Graduate School of Business, Knight Management Center, 655 Knight Way, Stanford, CA 94305, USA

HIGHLIGHTS

- People make self-serving attributions when responding to prosocial requests.
- But, morally self-serving attributions can backfire when challenged.
- People are forced into prosocial action if their excuse for refusal is removed.
- This also occurs if an external incentive (e.g., payment) is removed.

ARTICLE INFO

Article history:
Received 18 August 2016
Revised 2 November 2016
Accepted 3 November 2016
Available online 1 December 2016

Keywords:
Morality
Attributions
Decision making
Prosocial behavior

ABSTRACT

Two assumptions guide the current research. First, people's desire to see themselves as moral disposes them to make attributions that enhance or protect their moral self-image: When approached with a prosocial request, people are inclined to attribute their own noncompliance to external factors, while attributing their own compliance to internal factors. Second, these attributions can backfire when put to a material test. Studies 1 and 2 demonstrate that people who attribute their refusal of a prosocial request to an external factor (e.g., having an appointment), but then have that excuse removed, are more likely to engage in prosocial behavior than those who were never given an excuse to begin with. Study 3 shows that people view it as more morally reprehensible to no longer honor the acceptance of a prosocial request if an accompanying external incentive is removed than to refuse a request unaccompanied by an external incentive. Study 4 extends this finding and suggests that people who attribute the decision to behave prosocially to an internal factor despite the presence of an external incentive are more likely to continue to behave prosocially once the external incentive is removed than are those for whom no external incentive was ever offered. This research contributes to an understanding of the dynamics underlying the perpetuation of moral self-regard and suggests interventions to increase prosocial behavior.

The pervasive desire for people to see themselves as good and virtuous (e.g., Aquino & Reed, 2002; Blasi, 2004; Monin & Jordan, 2009) drives them to make self-serving attributions—blaming self-interested behavior on outside circumstances (Shalvi, Gino, Barkan, & Ayal, 2015) and taking credit for particularly moral behaviors. An important but unanswered question, however, is what happens when circumstances put those attributions to a material test. For example, suppose John justifies his decision not to donate to a particular charity with the claim that the money will go to overhead. How might he react when he finds out that the organization actually has low overhead costs? Will he simply shift to another comforting external attribution (e.g., the cause is not worthy) or will he feel trapped by his original attribution and now donate? Or, consider Jane, who has a self-interested reason to engage in a prosocial task (e.g., being paid), but insists that she would have volunteered regardless of external incentives. Now imagine circumstances were to change such that the self-interested benefit was removed (e.g., no longer being paid). Will she decide not to volunteer after all, or will she act consistently with her original claim that she would volunteer even if she was not paid?

John and Jane's situations above raise the question of how flexible the capacity for psychological self-protection is. People are motivated to make attributions regarding the reasons behind behavior (Bem, 1972; Kelley, 1973). In the case of their own behavior, people are motivated to, and amazingly resourceful at, making attributions that maximize self-regard (Kunda, 1990; Miller & Ross, 1975). Still, people's capacity to avoid unpleasant self-relevant conclusions is not infinite (Kunda, 1990). Both their denials and boasts must be plausible. Consequently, we argue that people have to "walk the talk" of their attributions, even if doing so is costly. In the context of prosocial requests, we propose that self-serving attributions, both for compliance and

* Corresponding author.
E-mail addresses: sclin1@stanford.edu (S.C. Lin), jjzlatev@stanford.edu (J.J. Zlatev), dtmiller@stanford.edu (D.T. Miller).

http://dx.doi.org/10.1016/j.jesp.2016.11.004
0022-1031/

noncompliance, can ironically backfire; people who confront altered circumstances that challenge their earlier attributions feel compelled to act in accordance with them.

In four studies, we test our prediction that self-serving attributions can backfire when challenged. In part I, we hypothesize that, when attributing their refusal of prosocial requests to external factors (e.g., "I have an appointment at the time"), people will be psychologically forced into prosocial action when those external factors are removed (e.g., the appointment is canceled). In part II, we hypothesize that people over-attribute their compliance with prosocial requests to intrinsic motivation and under-attribute it to external incentives and will be similarly psychologically forced into prosocial action when those incentives are removed.

1. Part I: backfiring of self-serving attributions during prosocial request refusal

Because people desire to maintain a moral self-image, they often feel uncomfortable engaging in self-other tradeoffs (Lin, Schaumberg, & Reich, 2016). Although individuals might prefer to take a self-interested action, they tend to want to do so without incurring moral self-reproach (Berman & Small, 2012). For this reason, people have developed justifications that allow themselves to engage in self-interested behavior while avoiding attributing that behavior to their own moral character (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996; Mazar, Amir, & Ariely, 2008; Pittarello, Leib, Gordon-Hecker, & Shalvi, 2015; Shalvi et al., 2015). For instance, people can take advantage of ambiguous moral wiggle room (Dana, Weber, & Kuang, 2007; Shalvi, Dana, Handgraaf, & De Dreu, 2011) or atone for their wrongdoings by engaging in prosocial or physically cleansing behaviors (Jordan, Mullen, & Murnighan, 2011; Shalvi et al., 2015; Zhong & Liljenquist, 2006). One more straightforward justification strategy relevant to the current research is attributing self-interested behavior to external forces (Snyder, Kleck, Strenta, & Mentzer, 1979).

The present research shows that offering external excuses for self-interested behavior is not cost-free and can backfire if circumstances remove the excuse. We argue that claims about why one did what one did in a particular situation can sow the seeds of a moral test (Miller & Monin, 2016). Consider the decision not to perform a particular prosocial behavior. The potential moral reproach from this decision can be mitigated by invoking a specific excuse (e.g., having an appointment at the time). Implicit in this excuse, however, is a moral claim (e.g., that one *would have* volunteered if possible) that heightens moral self-regard in the moment but sets a "trap" for the person if it is subsequently removed. To continue to refuse a prosocial request when the reason one initially gave for refusing it no longer applies (e.g., the appointment is canceled) would reveal not only that one has failed to live up to this heightened moral standard, but also that one is a hypocrite, which is a negative moral signal both to oneself and to others (Andrade & Ariely, 2009; Batson, Thompson, Seuferling, Whitney, & Strongman, 1999; Kreps & Monin, 2011). Accordingly, people cannot simply replace the former excuse with a new one, as it would undermine their earlier claim. Thus, we predict that when a person's excuse for not taking a prosocial action is removed, he or she will feel increased pressure to take that action.

We conducted Studies 1 and 2 to provide evidence that people who are given an excuse to refuse a prosocial request would be more likely to comply with the request when the excuse was removed than those who were never offered an excuse to begin with. We tested this using a scenario study (Study 1) followed by a study with a behavioral outcome (Study 2).

1.1. Study 1

Study 1 tested whether people who choose a self-interested action over a prosocial one when they have an external excuse for doing so would be more likely engage in the prosocial action when their original excuse was removed than those who never had such an excuse. We also sought to determine whether enhanced moral self-regard mediated this effect.

1.1.1. Methods

1.1.1.1. Participants. Given uncertainty about effect size, we adhered to the suggestion of using at least 50 participants per condition (Simmons, Nelson, & Simonsohn, 2013) and recruited 100 participants per condition, resulting in 200 participants from Amazon Mechanical Turk ($0.50 payment, M_{age} = 32.12, SD_{age} = 10.37, 114 men, 86 women). Data for this and subsequent studies were not analyzed until data collection was completed by Amazon Mechanical Turk (sometimes resulting in slightly larger samples than originally intended). We exclude no data for any study and report all the manipulations and measures used.

1.1.1.2. Procedure. Participants read a scenario that described how, in the spirit of the holidays, their town was offering food service at a local shelter, and that they were conflicted about whether to volunteer—they wanted to help the community, but it would also be a big hassle. In one condition, participants read that they realized they had an appointment at the same time. All participants then wrote their thoughts and feelings in free-response format. They then rated themselves on the nine traits (1 = *not at all characteristic of me*, 5 = *extremely characteristic of me*; $\alpha = 0.89$) previously shown to be viewed as indicative of moral character (Aquino & Reed, 2002), and filler items (creative, powerful, funny, athletic, disorganized, shy, neurotic) included to reduce demand effects. Afterwards, those who originally had an appointment were told that it was moved to a different day.

Everyone was then asked how obligated they felt to volunteer, how guilty they would feel if they did not volunteer, how good they would feel about themselves if they volunteered, how enjoyable volunteering would be (on 5-point scales, 1 = *not at all*, 5 = *extremely*), and how likely they would be to volunteer (0–100% slider). Finally, they were asked why they made the decision that they did in free-response format, and indicated gender and age.

1.1.2. Results

When people had an excuse not to volunteer, they rated themselves as higher on moral characteristics than when they did not have an excuse $t(198) = 3.87$, $p < 0.001$, see Table 1 for full statistics. In addition, of the filler variables, people reported feeling marginally more powerful and funny, and significantly more creative in the excuse condition than the control condition, and thus we control for these variables in mediation analyses below.

After the excuse was removed, those in the excuse condition expressed more obligation to volunteer than those in the no excuse condition, $t(198) = 6.48$, $p < 0.001$. Those in the excuse condition also reported, once it was removed, that they would feel more guilty if they did not volunteer, $t(198) = 6.48$, $p < 0.001$, and would find volunteering more enjoyable if they did volunteer, $t(198) = 3.39$, $p < 0.001$. They did not, however, report they would feel better about themselves if they did volunteer, $t(198) = 1.33$, $p = 0.185$. Finally, participants in the excuse condition indicated that they would be more likely to volunteer than those in the control condition, $t(198) = 5.48$, $p < 0.001$.

A serial mediation analysis (PROCESS Model 6, Hayes, 2013, see Fig. 1) revealed that those in the excuse condition rated themselves as more moral, leading to higher anticipated guilt if they did not volunteer, ultimately leading them to be more likely to volunteer 95% CI: [0.95, 4.44] (controlling for power, creativity and humor, 95% CI: [0.54, 3.70]).

1.1.3. Discussion

In Study 1, the availability of an external justification (excuse) for self-interested behavior (not volunteering) elevated participants'