



Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

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

Case Report



Attributional ambiguity reduces charitable giving by relaxing social norms[☆]

Fiona tho Pesch^{a,b,c,*}, Jason Dana^d

^a Institut Jean Nicod, Département d'études cognitives, Ecole normale supérieure, Université PSL, EHESS, CNRS, 75005 Paris, France

^b Max Planck Institute for Research on Collective Goods, Bonn, Germany

^c Department for Social and Personality Psychology, University of Wuppertal, Germany

^d Yale School of Management, Yale University, USA

ARTICLE INFO

Keywords:

Moral wiggle room
Attributional ambiguity
Prosocial behavior
Charitable giving
Social norms

ABSTRACT

A growing literature demonstrates reluctant giving: Many people who voluntarily give to charity no longer do so when they have an excuse not to give. The mechanisms of reluctance, however, remain unclear. Consistent with this literature, we found that injecting attributional ambiguity into a real charitable decision significantly reduces donations. Participants in our studies ($N = 2147$) faced a binary choice between options for distributing money between themselves and a charity, with one option giving more to a charity and the other leaving more for themselves. Borrowing from a classic attributional ambiguity paradigm, we manipulated whether the charity involved was the same for both options or different, giving participants the possible excuse of keeping more money due to preferring one charity over another. Participants indeed kept more for themselves when there were two different charities, regardless of which charity was associated with the more self-beneficial option, ostensibly revealing a hidden preference for selfishness. Using incentive compatible elicitions, we found no evidence that participants used the excuse of preferring one charity to another to justify their choices. Instead, we find that attributional ambiguity weakened perceptions that there is a norm against keeping more money in the task, both among decision makers and disinterested third parties. We conclude that attributional ambiguity lowers donations by relieving internalized social pressure to give.

1. Introduction

Charitable giving in the US is a growing multi-billion dollar business: In 2021, Americans donated an all-time record of \$485 billion (Giving USA Foundation, 2022). Why people donate is less clear. The question of motivational drivers of prosociality has occupied philosophers and scientists for centuries. Carefully constructed economic experiments show that people give to even anonymous others who cannot retaliate for not giving (e.g., Fehr & Fischbacher, 2002). As such, perhaps some people may give to charity for purely altruistic reasons (Batson & Shaw, 1991). Other researchers have argued that people may give because the act gives them a positive *warm glow* (Andreoni, 1990), meaning that giving makes them feel good about themselves. Alternatively, people may exhibit a form of inequality aversion (Fehr & Schmidt, 1999), in which

they experience negative reactions to an unequal distribution of resources, even if they would profit from it.

A burgeoning area of research in psychology and behavioral economics suggests that some giving, however, is *reluctant*. All of the motives described above lead people to want to give, feel good about giving, and welcome future opportunities to do so. Reluctant givers, on the other hand, voluntarily give when asked, but may experience negative feelings from doing so and avoid situations in which they will be asked to give (Cain, Dana, & Newman, 2014; Xu et al., 2023). For example, the same subjects who give in the aforementioned careful economic experiments keep more for themselves when they can avoid costless information about the impact of their choices on others' payoffs (Dana, Weber, & Kuang, 2007; Grossman & van der Weele, 2017). Thus, some people voluntarily give when they would otherwise prefer not to,

[☆] This paper has been recommended for acceptance by Dr. Kristin Laurin.

* Corresponding author at: Institut Jean Nicod, Département d'études cognitives, Ecole normale supérieure, Université PSL, EHESS, CNRS, 75005 Paris, France
E-mail address: fiona.tho.pesch@ens.psl.eu (F. tho Pesch).

<https://doi.org/10.1016/j.jesp.2023.104530>

Received 25 October 2022; Received in revised form 28 July 2023; Accepted 27 August 2023

Available online 13 September 2023

0022-1031/© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

apparently because they experience psychological costs associated with refusing to give (Bénabou & Tirole, 2006; Berman & Small, 2012; Lindsey, Yun, & Hill, 2007).

While it is well established that some giving is reluctant, it is less clear what mechanism is behind reluctant giving. The so-called self-image account proposes that people use excuses when available to behave in more self-beneficial ways while feeling they haven't violated their own standards of fairness. People generally want to establish and maintain a positive moral image of themselves (Aquino & Reed, 2002; Dunning, 2007; Monin & Jordan, 2009; Rachlin, 2002). But when one faces a prosocial request, such as an appeal for a charitable gift, this motivation stands in conflict with material self-interest. In such settings, refusing the request comes with psychological costs such as self-reproach (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996; Bandura, Caprara, Barbaranelli, Pastorelli, & Regalia, 2001; Higgins, 1997) and negative self-evaluation (Jordan, Leliveld, & Tenbrunsel, 2015; Rothmund & Baumert, 2014). Complying with the request avoids these psychological costs by paying the material costs of giving. It is reasonable, then, that people may seek excuses for why self-beneficial behavior does not actually violate their moral standards in such settings (Andreoni, Rao, & Trachtman, 2017; Bandura et al., 1996; DellaVigna, List, & Malmendier, 2012; Exley, 2016; Lin, Schaumberg, & Reich, 2016). This way, people can reap the benefits of the self-beneficial choice without paying the psychological costs of violating their own standards.

Another possible mechanism is social image. People don't like to violate what they see as social norms against selfishness (Andreoni & Bernheim, 2009; Bénabou & Tirole, 2006; Kimbrough & Vostroknutov, 2016; Krupka & Weber, 2013). This is apparently true even in anonymous experiments with no possibility for quid pro quo, where decision makers still care about what others think of their actions (Cain et al., 2014). Indeed, social norms often function in internalized ways, meaning that people follow social norms when not being observed by others at all (Bicchieri, 2005; Conte, Andrighetto, & Campenni, 2010). Rather than having a self-defined moral standard about how much to give, then, people may have an internalized social standard and give when not doing so would be perceived as violating a social norm.

Self-image and internalized social image are easily conflated in designs meant to demonstrate reluctant giving: If an experimental manipulation provides enough of an excuse to choose self-benefit without spoiling one's self-image, it probably provides enough excuse that (even imaginary) observers are unsure of the decision maker's motives. Thus, providing for self-excuses weakens perceived social norms against a self-beneficial choice.

We borrowed from a classic paradigm in social psychology (Snyder, Kleck, Strenta, & Mentzer, 1979) to investigate whether introducing attributional ambiguity into a charitable request will reduce giving, and why. In our experiments, participants faced a binary choice of how to allocate money to both themselves and a charity. One option was more self-beneficial (keeping more for themselves), while one was more charity-beneficial (giving more to charity). We manipulated whether the charities associated with each option were the same or different, and counterbalanced which charity was associated with the self-beneficial option. If participants were to systematically give less when the charities differed, they could be using attributional ambiguity as an excuse to give less while maintaining a positive self-image. That is, they could tell themselves that they actually prefer the charity associated with the higher payment to the self. To this end, we used incentivized elicitation mechanisms to see whether participants subsequently preferred the charity associated with the self-beneficial option. Alternatively, introducing ambiguity might have changed the perception of what is socially appropriate, making self-beneficial behavior less inappropriate. To this end, we also used incentivized elicitation methods to identify whether shared perceptions of the social norm changed under attributional ambiguity.

1.1. Reluctant giving

Though we see a great deal of prosociality around us, some of it is not as genuine as we would hope. Sometimes people help others or give to good causes when they would rather not, but give into the pressure such a request creates (Bursztyn & Jensen, 2017). As such, they also look for excuses for justifying self-interested behavior (Batson, Thompson, & Chen, 2002; Monin & Norton, 2003), reflecting a fundamental desire to be seen and to see oneself in a positive and moral light (e.g., Kruglanski, 1989; Kunda, 1990). In this intrapersonal conflict between a need to keep and a need to give, their *want-self* would rather keep its resources (i.e., time or money), while the *should-self* feels an obligation to give (Bazerman, Tenbrunsel, & Wade-Benzoni, 1998).

Studies on so-called *moral wiggle room* show that when the one-to-one link between the outcome of one's actions and one's intentions is broken, giving declines precipitously (Dana et al., 2007; Grossman & van der Weele, 2017). Participants playing simple economic games frequently give some part of an experimental endowment to other anonymous participants or to charities, even though the recipients will never learn their identity and cannot retaliate if they give nothing (Eckel & Grossman, 1996). This behavior would seem to reflect a preference for fair outcomes. Yet, levels of generosity significantly decline once participants have a way to maximize their own payoffs without revealing selfish motives. For example, many participants in these economic games avoid free information about the consequences for others of choosing a self-beneficial option (Dana et al., 2007; Feiler, 2014; Grossman, 2014; Grossman & van der Weele, 2017), thus allowing them to behave self-interestedly while avoiding the risk of feeling self-reproach or imagined reproach by anonymous others in the experiment. Likewise, it has been suggested that people strategically use the risk that charitable donations will be wasted as an excuse not to give (Exley, 2016). In field studies, people engage in costly avoidance of charitable requests when they could simply decline them. For example, people avoid the exits of a supermarket where Salvation Army bell ringers are asking for donations (Andreoni et al., 2017), or avoid being home if they know a charity solicitor will be coming (DellaVigna et al., 2012). Similarly, avoidance of prosocial requests occurs in the lab. Many experimental participants accept a smaller monetary payment (e.g., \$9) rather than a larger payment (e.g., \$10) if they will be asked, but not forced, to share the latter with an anonymous other participant (Dana, Cain, & Dawes, 2006; Lazear, Malmendier, & Weber, 2012; Lin et al., 2016). These studies all have in common that they allow people to maximize their own payoffs without appearing selfish, either to themselves or others, even when the others are anonymous or hypothetical.

1.2. Attributional ambiguity as an excuse

Classic studies of attributional ambiguity demonstrate experimental strategies for revealing participants' hidden, undesirable motives. When Snyder et al. (1979) asked participants to choose between two rooms to sit and watch the same movie, they were equally likely to sit next to a disabled person as a person without physical disabilities. However, when the movies in the two rooms were different, only 17% of the people sat in the room with the disabled person, even though the movies were counterbalanced across rooms. The authors thus concluded that some people truly desired to avoid the disabled, but only did so if this motive was not clearly revealed.

Correspondent inference theory (Jones & Davis, 1965) holds that the strength of inference one can make from observing someone's choices depends on the number of *noncommon effects* between the chosen and the forgone options. Noncommon effects are outcomes that are brought about by selecting one specific alternative, but not another. Decisions that differ only on one dimension (e.g., whether the person in the room has a disability) allow observers to attribute the agent's intention to that dimension. Having multiple noncommon effects (e.g., the people and the movies in the rooms) creates attributional ambiguity: It is not clear

which dimension drove the agent's choice. Attributional ambiguity thus allows one to choose according to their intrinsic preference with reduced concern about revealing undesirable motives to others. The effects of attributional ambiguity on behavior have been observed in studies of discrimination (Batson, Bolen, Cross, & Neuringer-Benefiel, 1986; Norton, Vandello, & Darley, 2004; Snyder et al., 1979) and willful ignorance (Woolley & Risen, 2021).

Attributional ambiguity in prosocial decision settings can thus be understood as an example of providing moral wiggle room. When people are asked to choose between donation amounts, the amount of money is the only noncommon effect. When introducing a second noncommon effect, for example different charities associated with the different donation amounts, it is not clear whether giving less reveals a selfish motive or a preference for a particular charity.

1.3. Identifying mechanisms

While the present work is inspired by the classic attributional ambiguity paradigm created by Snyder et al. (1979), those authors did not draw a strong conclusion about the channel through which attributional ambiguity worked. Indeed, there are several different channels through which attributional ambiguity could impact behavior according to correspondence inference theory (Jones & Davis, 1965), pointing towards an over-identification of the theory in this domain. The different channels that can be derived from correspondence inference theory vary in terms of who should be deceived about the true underlying motivation of certain behavior. Snyder et al. (1979) suggested that their participants may have sat with the confederate because they wanted to look good to themselves, to the experimenter, or to the confederate (Snyder et al., 1979).

Also in the literature on wiggle room, the underlying mechanisms driving the behavioral effect remain unclear. Here, a common argument is that giving drops off when wiggle room is introduced because people were hiding their true preference for selfish outcomes to maintain a positive self-image while still maximizing their own payoffs (Grossman & van der Weele, 2017; Matthey & Regner, 2015). That is, subjects have a moral self-standard about how much they should share, and moral wiggle room allows them to choose the self-beneficial option without feeling that they have violated this standard.

Self-image concerns have long been recognized as an important source of prosocial behavior (Barclay, 2004; Baumeister, 1998; Bem, 1972; Festinger, 1957; Fiske, 2009; Hardy & Van Vugt, 2006; Kawamura, Ohtsubo, & Kusumi, 2021; Konow, 2000; Willer, 2009). However, little direct evidence exists to show that self-image concerns drive reluctant giving. Woolley and Risen (2021) find that attributional ambiguity changes behavior both in public and private decision settings, and thus conclude that the effect is driven by self-image related factors. Self-image has also been hypothesized to drive the effect of moral wiggle room more broadly by several authors (Dana et al., 2007; Lazear et al., 2012; Matthey & Regner, 2015; Momsen & Ohndorf, 2020). Taking the example of using the risk that a charity might be ineffective as an excuse not to give, a self-image account would suggest that people actually believe that the reason they did not give was risk; i.e., that they used the excuses to themselves so that they could maintain a positive self-image (Bem, 1972; Goffman, 1959) without paying the monetary cost.

Other researchers have stressed the importance of social image concerns for reluctant giving (Andreoni & Bernheim, 2009; Grossman, 2015). People want to be seen as moral individuals by others (Blasi, 1980; Dunning, 2007; Monin & Jordan, 2009). As such, they generally follow social moral norms of fairness and prosociality (Andreoni & Bernheim, 2009; Gächter, Gerhards, & Nosenzo, 2017; Kimbrough & Vostroknutov, 2016; Krupka & Weber, 2013). Social norms can be seen as rules and standards of behavior within a group that proscribe selfish interests in favor of group interests by way of cooperation and prosociality (Ohtsuki & Iwasa, 2006; Thøgersen, 2008). In other words, observed or own behavior is compared to social standards in order to

judge its (in-) appropriateness, and potentially punish norm transgressions to uphold the social norm. Further, social norms often function in internalized ways, meaning that people also follow social norms when not being observed by others at all (Bicchieri, 2005; Conte et al., 2010). When social norms are internalized, people judge their own behavior by comparing it with socially defined standards, but sanctions or rewards are administered by the individual themselves in the form of experiencing guilt or pride (Kimbrough & Vostroknutov, 2016). Social norms have been implicated as a mechanism in experiments demonstrating reluctant giving (Bartling, Engl, & Weber, 2014; Conrads & Irlenbusch, 2013; Krupka & Weber, 2013). Indeed, some experimental participants whose giving would otherwise be influenced by social norms will refuse to receive information about the descriptive norm (Andersson, Erlandsson, & Västfjäll, 2022; Chan, Liao, Martin, & Wang, 2023). If excuses for not giving are readily apparent in a situation, then a failure to give does not clearly reveal a selfish motive and therefore, the normative pressure to give will be weakened.

To the extent that people could be responding to an internalized social standard, the interpretations of prior studies of reluctant giving are often confounded. Introducing wiggle room into the situation provides possible excuses why keeping more does not violate self-standards of prosociality. At the same time, by providing an apparent excuse for not giving other than selfishness, wiggle room may also make it less clear to onlookers what the decision maker's motives were. Thus, wiggle room relaxes the perception that a self-beneficial choice is socially inappropriate in a given situation, as observed behavior cannot be attributed to selfish motives. Returning to the example of risk as an excuse for not giving, if risk could be a reason that people do not give, then they do not necessarily reveal a selfish motive by not giving. That is, an observer would not be able to tell whether a person did not give because of selfish motives or because of the risk that the gift would be misappropriated. The norm of giving in this situation could thus be weakened. If people care about doing what is socially appropriate, they may give less when they have the excuse of risk, even if they do not use risk as an excuse to themselves to assuage self-image concerns. That is, even if people do not fool themselves into believing the reason they are not giving is risk, they may still give less through the channel of internalized norms: They are less likely to believe that giving is socially required in the situation.

The internalized social image account can thus explain many of the same findings as the self-image account. But it also accounts for findings that the self-image account does not. For example, the findings mentioned above on intentional avoidance of charitable requests (Andreoni et al., 2017; DellaVigna et al., 2012) and accepting smaller but non-shareable lab payments (Dana et al., 2006) involve knowingly avoiding charity. It is unclear how knowingly giving nothing squares with self-image, but charitable avoidance can be squared with internalized social image because avoidance prevents any observation of the charitable decision at all, and thus prevents there being a social image about which to be concerned.

This distinction between the self-image account and the internalized social image account might seem semantic at first blush. If social standards are truly internalized, and cause guilt when they are violated, then the social image account sounds essentially like a self-image account: People are still trying to avoid the guilt of having violated a standard they hold important. The psychological process, however, is different and leads to different predictions. On one account, the decision maker looks to rationalize that they have not violated their own standard (in the current studies, by convincing the self that the charity associated with the better self-payoff is more worthy). On the other account, the decision maker conforms to what they think is normative. By using incentivized methods of eliciting preferences among our charities and beliefs about social norms, we tease apart the possible pathways of self-image and internalized social image.

2. Overview of studies

In four studies, we investigated the effects of attributional ambiguity on charitable giving. We predicted that people would give less to charity when there is a reason not to give other than selfishness. We further investigated potential mechanisms.

In the first two studies, participants faced a binary choice between giving more or less to charity at the expense of their own pay in the study. We manipulated whether the donation went to the same charity in both options (Same Charity condition) or whether the charity differed between the options (Different Charities condition). We hypothesized that participants in the Different Charities condition were more likely to choose the option that maximized their own payoff because it would be ambiguous whether they were making that choice out of selfishness or out of a preference for the associated charity. In study 2, we replicated the task from study 1 in a larger sample, and explored potential mechanisms. Did participants fool themselves into thinking that they actually preferred the charity associated with the self-beneficial option (self-image account)? Or did they (correctly) perceive that under attributional ambiguity, there was less of a clear norm against the self-beneficial option (internalized social image account)? In our design, we investigated the self-image account by asking participants to vote for which charity should receive an additional donation of \$50 from the experimenters, with the money going to the charity that received the most votes. If participants in the Different Charities condition systematically voted for the charity that was linked to the self-beneficial option, it would be strong suggestive evidence that self-image concerns played a role in the effect of attributional ambiguity. To investigate the social image account, we elicited participants' perception of the prevailing social norm in study 2 using an incentive compatible elicitation method from Krupka and Weber (2013). If the self-beneficial option was perceived as less socially inappropriate under attributional ambiguity, this would be strong suggestive evidence that social image in the form of social norms played a decisive role in the effect of attributional ambiguity. Again, because participants in these experiments made their decisions privately, it would be specifically internalized social image at play.

Participants' ratings of social appropriateness in study 2, however, could be influenced by their prior behavior in the study. That is, if participants chose in a self-beneficial manner, they may be motivated to rate such behavior as more socially appropriate, perhaps more strongly motivated than by the monetary incentives in the experiment to correctly identify the norm. We addressed this possibility in study 3a by eliciting social norms for the Same and the Different Charities conditions from studies 1 and 2 from independent samples who had read about the study design but did not actually make allocation choices. We more directly corroborated the link between perceived norm changes and self-beneficial behavior in Study 3b by asking an independent sample to rate whether self-beneficial choices differed in their perceived selfishness across the Same and the Different Charities conditions.

3. Study 1

Study 1 examined the effect of introducing attributional ambiguity into a charity decision context where the charities we selected were rated as equally attractive during pretesting.¹ Following Snyder et al. (1979), we expected participants to be more likely to choose the self-beneficial option when there was attributional ambiguity (i.e., in the Different Charities condition), compared to when there was no attributional ambiguity (i.e., in the Same Charity condition). We furthermore examined self-image related elements of the mechanism, such as a change in charity preferences due to our manipulation.

3.1. Methods

Participants and design. Re-examining Snyder et al.'s data, we found that pooling their study 1 and study 2 yielded a large effect size of $V = 0.425$. Because replications generally lead to effect sizes that are smaller than the original (Open Science Collaboration, 2015) and our charitable giving context was slightly different, we recruited 240 subjects, which would yield a power above 0.8 to detect an effect half the size of the original. Eighteen participants were released from the study without making a choice after failing comprehension questions that ensured they understood the task, leaving us with 222 of subjects who took part in the study (see Materials in <https://osf.io/6jp9q> for comprehension questions). In all studies, including study 1, we recruited US-based participants from Amazon's Mechanical Turk with an approval rate of at least 98% and a minimum of 50 approved HITs. We excluded participants who had participated in any of our prior studies. The study took about 5 min. Participants received a flat fee of \$0.35 in addition to a bonus payment of \$0.40 to \$0.50, depending on their choice. Participants could also earn an additional \$0.10 by guessing the charity that would be rated most popular by a majority of other participants. The study was a between-subjects design, with 111 in each condition.

Procedure. In all conditions, participants first read the same general instructions, and answered two comprehension questions. Participants who failed the comprehension questions twice were then exited from the study and did not proceed to the decision stage. All other participants were then randomly assigned to one of the two conditions, and chose between two options that allocated money to themselves and a real charity (see Fig. 1). In the Same Charity condition, the donation went to the same charity in both options. In the Different Charities condition, the charities differed between the options. Participants read short descriptions of the charities used in study 1: No Lean Season enables labor mobility for the poorest in rural agricultural areas, while the END fund advances treatments to end neglected tropical diseases.² We counterbalanced the charities in the Different Charities condition, and randomly assigned participants to one of the charities in the Same Charity condition. Both the participants and the charities were paid according to participants' choices.

After the allocation decisions, all participants were asked to rate on a scale from 1 ("not at all") to 5 ("very much") how conflicted they felt about their decision, how satisfied or happy they were with their decision, and four questions gauging how trustworthy the charities were and how important charitable giving was. All participants then read the description of both charities, and answered two incentive compatible questions about the charity they preferred. First, they answered a *personal preference question* by voting for which charity should receive an additional donation of \$50 from the experimenters, with the money going to the charity that received the most votes. They then answered a *popularity question* in which they indicated which charity most other participants would think was more popular, receiving a bonus of \$0.10 if their answer was the most commonly given. For exploratory reasons, we asked participants to make a couple of hypothetical sharing decisions at the end of the experiment (see Appendix B2).

3.2. Results

Choices did not differ between the two counterbalanced Same Charity conditions, nor between the two Different Charities conditions ($ps > 0.472$), indicating that participants did not generally prefer one charity to the other. We therefore collapsed the data in each condition. As predicted, participants were less likely to choose the self-beneficial option in the Same Charity condition (36.0%), which had only one noncommon effect, than in the Different Charities condition (58.6%), which had attributional ambiguity, $\text{Chi}^2(1) = 11.29, p = .001, V = 0.23, 95\% \text{ CI } [0.10, 0.35]$ (see Table 1).

¹ For the results of our pre-study, see Appendix A.

² No Lean Season ceased operations sometime after we ran this experiment.

A	You: \$0.50 No Lean Season: \$0.20
B	You: \$0.40 END Fund: \$0.40

A	You: \$0.50 No Lean Season: \$0.20
B	You: \$0.40 No Lean Season: \$0.40

Fig. 1. Options in the Different Charities condition (left) and the Same Charity condition (right).

For the personal preference question, 62% of all participants voted for the END Fund to receive the additional donation, while 60% picked the END Fund for the popularity question. To answer whether the effect of the Same vs. Different Charities manipulation was consistent with self-image concerns, we compared the personal preference question across the two counterbalanced versions of the Different Charities condition. Participants were not significantly more likely to vote for the END Fund when it was associated with the self-beneficial option (62.5%) than when it was associated with the charity-beneficial option (58.2%), $\chi^2(1) = 0.22, p = .642, V = 0.04, 95\% \text{ CI } [-0.14, 0.23]$. Similarly, there was no difference in the popularity question between the two counterbalanced Different Charities conditions (58.9% vs. 54.5% guessing that others preferred the END Fund), $\chi^2(1) = 0.22, p = .641, V = 0.04, 95\% \text{ CI } [-0.14, 0.23]$. When looking at how trustworthy participants perceived the two charities in the self-report measures, we did not find any differences between the two counterbalanced Different Charities conditions, all $ps > 0.100$ (for more analyses on these self-report measures, see Appendix B1).

3.3. Discussion

Introducing attributional ambiguity into a charitable decision context increased choice of the self-beneficial option from 36.0% to 58.6%. We thus conceptually replicate Snyder et al. (1979) in the domain of charitable giving, revealing a hidden preference for the self-beneficial option. We did not find direct evidence of self-image concerns by way of people indicating that they chose according to the second noncommon effect: Participants were not more likely to report a personal preference for the charity that matched the self-beneficial option than the charity that matched the charity-beneficial option. In other words, our participants did not appear to blame their self-beneficial choices on a preference for a specific charity. Study 2 replicated these findings and investigated whether attributional ambiguity weakens the social norm against selfishness and whether social norms mediate the effect of attributional ambiguity on giving.

4. Study 2

Study 2 sought to replicate the effect of attributional ambiguity on charitable giving and to investigate its effect on perceived social norms. We used a method that elicits true beliefs about norms through incentivized choices (Krupka & Weber, 2013). Specifically, participants indicated the social appropriateness of certain behavioral responses,

Table 1

Participants were more likely to choose the self-beneficial option in the Different Charities condition.

	Same Charity	Different Charities	
self-beneficial	40 (36.0%)	65 (58.6%)	105
charity-beneficial	71 (64.0%)	46 (41.4%)	117
	111	111	222

while being incentivized for picking the option that was chosen by most other participants. Participants thus had incentives to state their true belief about the social appropriateness of each option.

4.1. Methods

Because we added social norms as a factor to investigate and wished to distinguish it from self-image explanations, we preregistered a highly-powered study seeking 750 participants per condition including study design, hypotheses and analysis plan. Preregistration materials are available at <https://osf.io/6j9p9q>. After 8 participants failed the comprehension questions and were dismissed before beginning the study, we were left with 1492 participants from Amazon Mechanical Turk (54% female). A sensitivity analysis revealed that this sample size yielded 80% power to detect an effect size of $V = 0.073$. We replicated the set-up used in study 1, adding the social norm elicitation method of Krupka and Weber (2013) to the post-experimental questionnaire. On a scale from 1 (“Very socially inappropriate”) to 4 (“Very socially appropriate”), participants were asked to indicate the social appropriateness of each option in their respective experimental conditions (i.e., social norm). Participants were informed that one of the two options was selected randomly to determine a bonus payment of \$0.10 if the participant’s response was the same as the most common response. We also asked participants to indicate how appropriate they personally found each of the two behavioral options (i.e., personal norm), ranging from 1 (“Very inappropriate”) to 4 (“Very appropriate”).

4.2. Results

As in study 1, we first tested whether participants significantly favored one charity over another, and found no significant differences across counterbalanced conditions, $ps > 0.322$. The main effect of our Different Charities manipulation was smaller than in study 1, but statistically significant, $\chi^2(1) = 25.13, p < .001, V = 0.13, 95\% \text{ CI } [0.08, 0.18]$: In the Same Charity condition, about 32.2% of participants chose the self-beneficial option, increasing to 44.8% in the Different Charities condition (see Table 2).

As in study 1, we did not observe a subsequent effect of our manipulation on participants’ personal preferences over the two charities in the two counterbalanced versions of the Different Charities condition, $\chi^2(1) = 0.07, p = .799, V = -0.01, 95\% \text{ CI } [-0.08, 0.06]$. However, these participants thought that the charity linked to the self-beneficial option would be more popular, $\chi^2(1) = 7.77, p = .005, V = 0.10$,

Table 2

Participants in the Different Charities condition were more likely to select the self-beneficial option.

	Same Charity	Different Charities	
self-beneficial	240 (32.2%)	335 (44.8%)	575
charity-beneficial	505 (67.8%)	412 (55.2%)	917
	745	747	1492

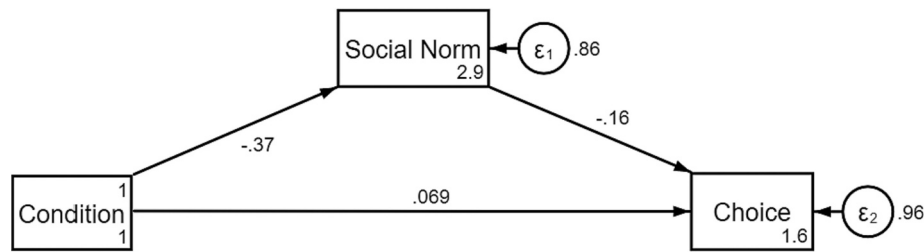


Fig. 2. Path diagram (with standardized coefficients) displaying the mediation of attributional ambiguity (Condition: 0 = Different Charities, 1 = Same Charity) on prosocial behavior (0 = self-beneficial choice, 1 = charity-beneficial choice) through social norms concerning behavior.

95% CI [0.03, 0.17]. A similar pattern can be seen in the two counter-balanced versions of the Same Charity condition: Participants thought that the charity they just donated to would be more popular amongst others, $\chi^2(1) = 5.36, p = .021, V = 0.09, 95\% \text{ CI } [0.01, 0.16]$, but this was not significantly related to their personal preference, $\chi^2(1) = 2.48, p = .115, V = 0.06, 95\% \text{ CI } [-0.01, 0.13]$.

Our manipulation also had a significant effect on the perceptions of social norms: People perceived choosing the self-interested option to be more socially appropriate in the Different Charities condition ($M = 2.97, SD = 0.98$) than in the Same Charity condition ($M = 2.21, SD = 0.92$), $t(1490) = 15.51, p < .001, d = 0.80, 95\% \text{ CI } [0.70, 0.91]$. The social appropriateness of behaving prosocially decreased slightly, but significantly from $M = 3.80 (SD = 0.54)$ in the Same Charity condition to $M = 3.66 (SD = 0.65)$ in the Different Charities condition, $t(1490) = -4.43, p < .001, d = -0.23, 95\% \text{ CI } [-0.33, -0.13]$. We conducted a mediation analysis to see if perceived social norms regarding selfishness mediated the effect of attributional ambiguity on giving (see Fig. 2). The indirect effect of attributional ambiguity through perceived social norms was significant, $\beta = 0.059, \text{ Sobel } Z = 5.56, p < .001$, mediating about 47% of the effect. Personal norms follow the same pattern (see Appendix B3 for a discussion of the role of personal norms).

4.3. Discussion

In study 2, we replicated the behavioral effect of attributional ambiguity on charitable giving. We also shed light on the potential mechanism behind it: While participants did not seem to change their personal preference for one of the charities, which would be consistent with self-image concerns, they did perceive a change in the social norm: Self-beneficial behavior was expected to be seen by others as more socially appropriate in the Different Charities condition. As such, choosing the self-beneficial option might have a less negative impact on the agent's social image. The possible influence of social norms is interesting because our setup, like many investigating reluctant giving, involved only private behavior; participants do not observe the behavior of other participants. Therefore, our results are most consistent with a concern for abiding by internalized social norms (Bicchieri, 2005; Conte et al., 2010). Our mediation results suggest that attributional ambiguity increased self-beneficial behavior by making it more socially appropriate, but causality cannot be strictly inferred from our experimental setup. Specifically, the reported social norm ratings could reflect a shared belief that most participants also want to rationalize their self-beneficial choices after the fact. Study 3a thus examined whether attributional ambiguity changed the perceived social norm using a sample that did not make the charity decision, and thus had no motivations to rationalize their own choices. Study 3b investigated the relation of this change in norm to the perception of behavior as selfish.

5. Study 3a & b

5.1. Methods

We recruited 485 participants from Amazon Mechanical Turk for both studies 3a and b, yielding a power of 0.8 to detect a small effect of $d = 0.23$. After some participants were dismissed for failing comprehension checks, we were left with a final sample size of $N = 433$ in study 3a, and $N = 460$ in study 3b. The first part of the instructions was identical to studies 1 and 2, except that the instructions talk about another "person X" who was to perform the decision stage. After passing the comprehension questions, participants in study 3a were asked to rate the social appropriateness of each choice as described in study 2 (i.e., social norms; Krupka & Weber, 2013). This included an extra \$0.50 bonus payment if a participant's response was the most frequent response given by other participants. We then asked participants to indicate how appropriate they personally found each of the two behavioral options (i.e., personal norms), ranging from 1 ("Very inappropriate") to 4 ("Very appropriate"). In study 3b, we asked participants about how selfish each behavioral option would be on a scale from 1 ("Not selfish at all") to 5 ("Very selfish"). Participants then answered three questions about how much they would feel guilty/regret/have a bad conscience if choosing option A on a scale from 1 ("Not at all") to 5 ("Very much"). Study 3b did not involve incentivized choices because we asked for participant's own personal ratings of how selfish each option was. Participants in study 3b simply received an extra \$0.20 if they passed the comprehension questions. In both studies 3a and 3b, we manipulated between-subject whether participants were instructed about the Same or the Different Charities condition.

5.2. Results

In study 3a, participants in the Different Charities condition rated the self-beneficial choice to be more socially appropriate ($M = 2.76, SD = 0.89$) than participants in the Same Charity condition ($M = 2.42, SD = 0.88$), $t(431) = 4.04, p < .001, d = 0.388, 95\% \text{ CI } [0.20, 0.58]$. At the same time, charity-beneficial behavior (i.e., choosing option B) was perceived as slightly but significantly less socially appropriate in the Different Charities condition ($M = 3.63, SD = 0.66$) than the Same Charity condition ($M = 3.78, SD = 0.51$), $t(431) = -2.66; p = .004, d = 0.26, 95\% \text{ CI } [-0.44, -0.07]$. Again, personal norms follow the same pattern (see Appendix B3 for a discussion of the role of personal norms).

In study 3b, participants in the Different Charities condition rated the self-beneficial choice as less selfish ($M = 3.20, SD = 1.20$) than participants in the Same Charity condition ($M = 3.88, SD = 1.53$), $t(458) = -5.32, p < .001, d = -0.50, 95\% \text{ CI } [-0.68, -0.31]$. When asking participants how they would anticipate feeling if they had chosen option A, participants in the Different Charities condition anticipate feeling less

guilty ($M = 2.83$, $SD = 1.44$) than participants in the Same Charity condition ($M = 3.06$, $SD = 1.40$), $t(458) = -1.68$, $p = .047$, $d = -0.16$, 95% CI [-0.34, 0.03]. Similarly, participants in the Different Charities condition reported lower levels of anticipated bad conscience, $t(458) = -2.36$, $p = .009$, $d = -0.22$, 95% CI [-0.40, -0.04], and a trending effect in this direction on anticipated regret, $t(458) = -1.54$, $p = .06$, $d = -0.14$, 95% CI [-0.33, 0.04] if they chose the self-beneficial option than did participants in the Same Charity condition.

5.3. Discussion

Participants who did not make charitable decisions, and thus had no motivation to justify their own behavior, still perceived that choosing option A (the self-beneficial choice) was not as socially inappropriate when attributional ambiguity was introduced. The effect size, however, was smaller for these participants than it was in study 2, possibly indicating additional post-decision rationalization for participants who choose selfishly. In isolation, this result could mean either that self-beneficial behavior is seen as less inappropriate under attributional ambiguity, or that choosing the self-beneficial option is seen as a less selfish behavior. Our results from study 3b speak for the latter: Choosing option A is seen as less selfish, indicating that attributional ambiguity changes the way we evaluate behavior in terms of its morality.

6. General discussion

Using a classic paradigm that builds on correspondent inference theory (Jones & Davis, 1965; Snyder et al., 1979), we show that introducing attributional ambiguity into a charitable request significantly reduces giving. Specifically, our participants chose between two options, with one option giving more to a charity and the other leaving more for themselves. People chose the self-beneficial option more often when the options involved two different charities than when both options involved the same charity. Our findings conceptually replicate the findings of Snyder et al. (1979) in the domain of charitable giving. Our results are thus also in line with correspondent inference theory (Jones & Davis, 1965): By introducing a second noncommon effect to the decision setting (i.e., different charities associated with the self-beneficial and charity-beneficial options), a potential observer cannot draw clear dispositional inferences about the decision maker from observing one single decision. Only a large number of observations across conditions could allow us to identify what looks like a hidden motive of selfishness. We thus identified attributional ambiguity as one more form of moral wiggle room (Dana et al., 2007).

Correspondent inference theory, however, is over-identified in our context because there are multiple channels through which the theory could be confirmed. Specifically, having two different charities could provide our subjects with the convenient excuse of preferring the charity associated with the self-beneficial option, which they may fool themselves into believing is true (the self-image account). Alternatively, having two different charities could weaken the social norm against choosing the self-beneficial option because an onlooker could not infer selfish motives from such a choice (the internalized social image account). Crucially, both accounts would be in line with correspondent inference theory.

Our data supports the idea that the observed behavioral effect is driven by social, rather than self-image concerns. We found no support for the idea that people fooled themselves into thinking they preferred the charity linked to the self-beneficial option using both incentivized decisions and self-reports, which would have been in line with a self-image account. Rather, people seem to respond to their own perceptions of what others would think of their behavior, supporting an internalized social image account: Introducing attributional ambiguity reduced the perception that generous behavior was socially expected and, even though choices were not observed by others, relieved the perceived normative pressure to give. In line with correspondence

inference theory (Jones & Davis, 1965), the act of choosing the self-beneficial option is perceived as less selfish, potentially because the observed behavior cannot be clearly attributed to selfish motives due to the second noncommon effect in the setup (i.e., the different charities). We also observed this change in the perception of the social norm and selfish motives in separate samples of participants who did not make a distribution decision themselves. Self-beneficial choices were seen as less socially inappropriate and less selfish under attributional ambiguity by neutral third parties. Attributional ambiguity weakens signals of selfishness that behavior may carry, which results in less pressure coming from social norms. It appears sufficient that outside observers cannot clearly infer one's type with respect to prosociality, even though the decision makers themselves may be aware that they are choosing the self-beneficial option out of selfish motives.

That our participants changed their behavior according to changing norms, despite their decisions being unobservable is evidence of internalized norm-following (Bicchieri, 2005; Conte et al., 2010). It seems that people compare their own behavior to socially defined standards rather than their own private standards when they judge the morality of their own behavior. We stress that the decisions our participants made are not actually observed by other parties, including the charities. It is apparently an internalized form of social image that drives the effect of attributional ambiguity on sharing decisions. Our account is further supported by our finding on anticipated guilt: When social norms are internalized, rewards and sanctions are not administered by others, but by the individuals themselves in the form of guilt and pride (Kimbrough & Vostroknutov, 2016). In Study 3b, we observe that people anticipate experiencing less guilt for choosing the self-beneficial option under attributional ambiguity.

We argue that this over-identification problem is more general to studies of attributional ambiguity and to the literature on reluctant giving. For example, Snyder et al. (1979) definitively showed that adding attributional ambiguity led people to avoid a handicapped confederate, revealing that their participants had a hidden motive for doing so. They conceded, however, that it was not clear from whom the motive was being hidden: participants may have wanted to appear unbiased to themselves, to the experimenter, or to the experimental confederate. Through the use of careful elicitation methods, particularly norm elicitation methods, we conclude that our evidence is more in line with the concerns for appearance to others. Specifically in our task, as in many laboratory tasks that examine giving to others or to charity, nobody actually observed our participants; their concerns about social norms were internalized.

The reluctant altruism literature has also shown that mechanisms similar to attributional ambiguity reduce giving, revealing that participants have hidden motives to be more self-interested. But as in Snyder et al., it is unclear in much of this literature from whom motives are being hidden. Several authors have argued that because the decision maker's behavior is not observed by others, it must be self-image driving the effect of moral wiggle room (Grossman & van der Weele, 2017; Matthey & Regner, 2015; Woolley & Risen, 2021). That is, reluctant giving occurs when keeping would reveal unacceptable motives to the self, but the presence of excuses allows reluctant gives to keep more without revealing to themselves that they have violated their own moral standards. An internalized social norm is often, however, a competing explanation. If an experimental manipulation provides enough of an excuse to hide motives from the self, it will usually also hide motives as seen from the eyes of others. As such, these experimental excuses weaken perceived social norms against self-beneficial behavior. Returning to the example of Exley (2016), it appears that people use the risk that charitable contributions will not be used as intended as an excuse not to give. But the presence of risk also weakens the signal that a self-beneficial selfish choice sends to others about the agent's underlying motives. This means that the socially defined standards to which people compare their behavior have changed, so that self-beneficial behavior becomes more socially appropriate.

The distinction between the self-image account and the internalized social image account might appear pedantic. Both the internalized social image account, as well as the self-image account involve an internal evaluation of one's own behavior regarding some important standards. As such, they could be argued to both be *self-image* accounts. The underlying processes, however, are psychologically different in important ways. While a self-image account would assume that one's own behavior is compared to one's own moral standards, the internalized social image proposes that socially defined standards are the decisive factor. The internalized social image account can explain findings on reluctant altruism that the self-image account apparently cannot. For example, laboratory findings on dictator exit (Dana et al., 2006) and field studies showing intentional avoidance of charitable requests (Andreoni et al., 2017; DellaVigna et al., 2012) involve knowingly avoiding charity. Further, the dictator exit results show that many of those who knowingly avoid charity were sincerely going to give (Dana et al., 2006). It is unclear how knowingly giving nothing squares with self-image. Charitable avoidance can be squared with internalized social image, however, because avoidance prevents any observation of the charitable decision at all, and thus prevents there being a social image about which to be concerned.

Because (internalized) social standards are defined by judging behavior from an outside observer's perspective, attributional ambiguity can change internalized social standards independently from self-standards. To our knowledge, our work is the first one of its kind to tease apart the standards to which people comply when exploiting moral wiggle room, clearly speaking in favor of socially defined standards over private standards. Future research should explore whether the effect of attributional ambiguity on behavior is even more pronounced when sharing decisions are made in the presence of others.

So, why do people give to charity? While there are many motivations, it appears that people often give reluctantly in the presence of a request. Indeed, it has been suggested that as much as 50% of giving in lab and field experiments is done reluctantly and givers would have preferred to avoid the request or have an excuse not to give (Cain et al., 2014). Our results suggest that internalized social image, rather than self-image concerns, are a key driver. Social norms about what is appropriate in this arena are fragile. The introduction of attributional ambiguity reduces shared notions of what is socially appropriate, and people give less as a consequence.

Our results have multiple implications for charitable fundraising. First, we underscore the importance of understanding mechanisms behind reluctant giving. Even when choosing privately, giving decisions may be driven by internalized perceptions of what is socially appropriate. This motivation can be leveraged in charitable requests, for instance, by highlighting others' gifts or presenting favorable information about the frequency of giving. These implications are different than those of a self-image account. The self-image account, in our paradigm, involved fooling oneself about charitable preferences and would suggest providing comparative information about the worthiness of the charities. Second, we show the importance of understanding reluctant giving

Appendix A. Pre-study - Attractiveness of Charities

We sampled 50 participants from Amazon Mechanical Turk. The study took about 6 min, and participants received a payment of \$0.80. Participants were asked to rate six different charities on 5 items each on a scale from 1 ("Not at all") to 5 ("Very much"). The order of the charities were randomized between participants. For a list of the items, see Table 1. For a list of the charities and their respective description, see Table 2. In Table 3, you see the mean scores for each item and each charity. Table 4a, 4b and 4c presents different correlations between the charities. We decided to use The END Fund and No Lean Season for our main study. We picked these charities, as the correlations between the evaluations were high, and the absolute scores were relatively average compared to the other charities.

more generally as it relates to charitable requests. The same people who might give in response to normative information may also avoid receiving knowledge about descriptive norms so that they can give less (Andersson et al., 2022). It is thus important that requesters not unintentionally inject ambiguity into the charitable request situation. Possible excuses for not giving will relieve the pressure to give from internalized social norms. These excuses, as we show, can be subtle and could occur from something as seemingly innocuous as providing more variety of options. Finally, we continue a vein of research into the impact of charitable requests on givers themselves (Andreoni et al., 2017; DellaVigna et al., 2012). If some givers are reluctant and would prefer excuses not to give, then it is likely that the welfare of would-be givers is lowered by high-pressure requests, perhaps particularly those that use social pressure. As DellaVigna et al. (2012) point out, aggressive charitable marketing could have a negative overall impact if the charity is not highly efficient.

7. Conclusion

Attributional ambiguity allows people to maximize their own payoffs by making self-beneficial behavior less socially inappropriate. People do not try to fool themselves into thinking that they actually prefer the charity that is attached to the self-beneficial choice, which would have been in line with a self-image account. However, people selfishly benefit from the ambiguity of what motivated their choice because of the reduction in social expectations to give, which speaks for an internalized social image account. Future research should investigate the causality in the link between a change in social norms and subsequent behavior.

Open practices

We report how we determined our sample size, all data exclusions, and all manipulations in the study. All studies were granted exemption by the University's Human Subjects Committee (protocol number: 2000020511). All data, analysis code, measures, and research materials for all three studies, and the two studies reported in the Appendix are available at <https://osf.io/6jp9q>. Data were analyzed using STATA, version BE 17.0. Design and analysis were not pre-registered for study 1 and 3, but for the main effect of study 2 (see <https://osf.io/6jp9q>).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

For all data, analysis, preregistration and materials see <https://osf.io/6jp9q>.

Table 1
Items asked for each of the charities.

Variable name	Question wording
likely trust	Please imagine you would want to donate to a charity. How likely would you be to donate to this charity? How trustworthy do you find this charity?
better	To which extent do you think this charity makes the world a bit better?
important	How important is the cause of this charity?
avail	If you were asked to name charities that easily come to your mind, how likely would you be to think of this charity?

Table 2
Evaluated charities and their descriptions.

Charity name	Description
GiveDirectly	... is a charity operating mostly in Africa that helps families living in extreme poverty by simply sending them money with no strings attached. GiveDirectly describes their approach as simple and transparent, and also as relying on evidence. They work with independent researchers to demonstrate the effectiveness of giving cash directly.
The END Fund	... is recommended specifically for its deworming programs. Deworming is cheap and easy to administer, but the WHO estimates that at least 218 million people required preventive treatment for schistosomiasis in 2015, but only 66.5 million were treated. Its work is primarily focused on treating children, and based in Africa where the burden of disease caused by worm infections is the highest.
No Lean Season	... reduces the effects of seasonal poverty in rural agricultural areas by helping people travel for work. In rural areas, the season between planting and harvesting is known as the "lean season" due to low income and hunger. The program gives a travel subsidy of \$20 to very poor rural laborers so they can send a family member to a nearby city to find a job during this time. Migrant households register the equivalent of an extra meal per person / per day during the lean season.
The Against Malaria Foundation	... fights malaria by distributing insecticide-treated mosquito nets. Malaria killed around 438,000 people in 2015, including an estimated 306,000 children. There are about 200 million cases of the disease every year. Insecticide treated bednets are one of the most effective ways to prevent transmission of malaria and have averted about 450 million cases from 2000 to 2015. The Against Malaria Foundation (AMF) fights malaria by funding and tracking/monitoring insecticide-treated mosquito nets in Sub-Saharan Africa.
Make-A-Wish Foundation	... is a charity that creates life-changing wishes for children with a critical illness. A wish experience can be a game-changer for a child going through so much and impacts not only the kids but everyone involved - parents, volunteers, donors, sponsors, medical professionals and communities. Parents might finally feel like they can be optimistic. And still others might realize all they have to offer the world through volunteer work or philanthropy.
Rainbow Trust	... supports families who have a child with a life threatening or terminal illness and need the bespoke support we offer. Many of these children and their families are able to cope or are not in a 'crisis situation'. However, thousands of families have to face the very real possibility that their child may die and struggle to cope on a day to day basis. Our Family Support Workers provide a lifeline to these families and children. We support the whole family including parents, carers, the unwell child, brothers, sisters and grandparents. They bring support and help to families who so desperately need it at home, in hospital and in the community.

Table 3
Mean scores on each item for each charity.

	Against Malaria	END Fund	No Lean Season	Give Directly	Make-A-Wish	Rainbow Trust
likely	3.48	3.29	3.25	3.35	3.75	3.52
better	4	3.83	3.81	3.73	4.16	4.16
likely+better	3.74	3.56	3.53	3.54	3.93	3.82
trust	3.79	3.69	3.48	3.5	4.04	3.73
important	4.13	3.75	3.63	3.87	4	4.08
all eval.	3.85	3.64	3.54	3.61	3.98	3.86
available	3.10	2.56	2.85	3	4.31	3

Table 4a
Correlations between the charity evaluations, including all evaluative items.

	Against Malaria	END Fund	No Lean Season	Give Directly	Make-A-Wish	Rainbow Trust
Against Malaria	1					
END Fund	0.66	1				
No Lean Season	0.51	0.61	1			
GiveDirectly	0.58	0.45	0.53	1		
Make-A-Wish	0.34	0.41	0.26	0.36	1	
Rainbow Trust	0.59	0.48	0.57	0.66	0.52	1

Table 4b

Correlations between the charity evaluations, only including the two items Likely and Better.

	Against Malaria	END Fund	No Lean Season	Give Directly	Make-A-Wish	Rainbow Trust
Against Malaria	1					
END Fund	0.71	1				
No Lean Season	0.58	0.63	1			
GiveDirectly	0.69	0.51	0.58	1		
Make-A-Wish	0.43	0.44	0.35	0.55	1	
Rainbow Trust	0.63	0.48	0.62	0.72	0.58	1

Table 4c

Correlations between Avail.

	Against Malaria	END Fund	No Lean Season	Give Directly	Make-A-Wish	Rainbow Trust
Against Malaria	1					
END Fund	0.69	1				
No Lean Season	0.68	0.78	1			
GiveDirectly	0.61	0.61	0.61	1		
Make-A-Wish	0.05	-0.19	-0.16	0.09	1	
Rainbow Trust	0.52	0.67	0.72	0.66	-0.14	1

Appendix B. Additional analyses*Appendix B1: Additional analyses - Self reports (Study 1)*

Our experimental manipulation did not affect how satisfied or happy people were with their choice, all p s > 0.101. It also did not influence how important people deem charitable giving in general, $t(220) = 1.177$, $p = .240$, $d = 0.158$, 95% CI [-0.11, 0.42]. We do see a significant difference in how responsible people feel to donate to good causes: Participants in the Same Charity condition feel less responsible compared to people in the Different Charities condition, $t(220) = 2.068$, $p = .040$, $d = 0.278$, 95% CI [0.01, 0.54]. Also when it comes to how conflicted people felt about making the decision between the two options, we see differences between our two experimental conditions: participants in the Same Charity condition felt less conflicted compared to people in the Different Charities condition, $t(220) = 3.407$, $p < .001$, $d = 0.457$, 95% CI [0.19, 0.72].

Appendix B2: Additional analyses - Hypothetical allocation decisions (Study 1)

In order to explore whether we can identify single participants who are susceptible to our experimental manipulation, as well as to investigate whether participants will stick to the charity they donated to in the incentivized choice, we added hypothetical scenarios at the end of the experiment. All participants answered to two additional donation scenarios. In the first hypothetical decision, all participants faced one of the counterbalancing versions of the Different Charities condition they have not been faced with in the incentivized decision. In the second hypothetical decision, participants who faced a Different Charities condition in the incentivized choice setting were faced with the Same Charity decision setup. Participants who faced the Same Charity condition faced the other counterbalanced condition of the Different Charities decision setup, so that all participants faced both counterbalanced versions of the Different Charities setup and one version of the Same Charity setup throughout the experiment.

Our results indicated that participants who chose selfishly in the Different Charities condition were also more likely to choose the self-beneficial option in the first hypothetical scenario, when the other charity was linked to the self-beneficial option compared to the incentivized scenario, $r = 0.615$, $p < .001$.

Analyzing the second hypothetical scenario that participants facing the Same Charity setup are less likely to choose the charity-beneficial option (49.5%) compared to participants facing the Different Charities setup (63.1%), $\chi^2(1) = 4.120$, $p = .042$, $V = -0.136$, 95% CI [-0.27, -0.01].

Both hypothetical decisions are significantly correlated with the incentivized decision, and also with one another across all participants, with correlations ranging from 0.55 to 0.61.

These results suggest that we find strong carry-over effects within this experiment. All choices are strongly correlated with one another. In the second hypothetical decision, we even observe a reversal of the pattern of the incentivized choice: Now participants in the Same Charity condition are more likely to choose the self-beneficial option. This is most likely because these participants have been exposed to the Different Charities condition in the incentivized part of the experiment and have been more likely to choose selfishly there. Thus, this pattern is most likely explained by the strong carry-over effects. Participants simply stick to their initial decision of choosing the self-beneficial or the charity-beneficial option. This can be seen as tentative evidence that the effect of attributional ambiguity on social decision-making can carry over and generalize to settings in which this ambiguity is not given anymore.

Appendix B3: Additional analyses - Personal norms (Study 2 & 3a)

Though social norms undoubtedly are important for social behavior, recent studies suggest that personal norms are even more predictive of behavior (Bašić & Verrina, 2021). Prior empirical work suggests that while social and personal norms are correlated, but distinct, uniquely explaining variation in behavior (Bašić & Verrina, 2021). Thus, we also asked participants for their personal norms concerning the appropriateness of choosing option A or option B, both in study 2 and study 3a. In both studies, personal and social norms regarding selfish behavior are correlated (Study 2: $r = 0.70$; Study 3a: $r = 0.69$). In study 2, we find that behaving selfishly is seen as less inappropriate when attributional ambiguity is introduced, $t(1490) = 12.192$, $p < .001$, $d = 0.631$, 95% CI [0.53, 0.74]. We conducted a mediation analysis with personal norms regarding selfish behavior as a mediator

(see Fig. B3.1). The indirect effect of our experimental manipulation through perceived social norms regarding choosing the self-beneficial option on prosocial behavior was significant, $\beta = 0.110$, Sobel $Z = 9.46$, $p < .001$. Importantly, when plugging personal norms into the mediation model instead of social norms, the mediation becomes stronger. While the indirect effect through social norms explains about 47% of the variance, this increases to 87% when investigating personal norms. Thus, we conceptually replicated the effect of Bašić and Verrina (2021) that personal norms are a stronger predictor for behavior than social norms.

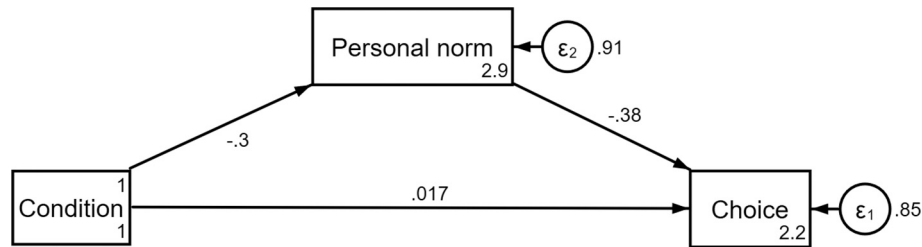


Fig. B3.1. Path diagram (with standardized coefficients) displaying the mediation of attributional ambiguity (Condition: 0 = Different Charities, 1 = Same Charity) on prosocial behavior (0 = self-beneficial choice, 1 = charity-beneficial choice) through personal norms concerning selfish behavior.

In study 3a, we see that not only do the social norms change in an independent sample who does not make this decision of choosing A or B themselves, but also the personal norms change. We see a similar effect in people's personal norms: again, selfish behavior seems more appropriate in the Different ($M = 2.82$, $SD = 0.90$) than in the Same Charity condition ($M = 2.60$, $SD = 1.05$), $t(431) = 2.33$; $p = .010$, $d = 0.223$, 95% CI [0.03, 0.41]. This suggests that also personal norms change due to the introduction of a second noncommon effect which diffuses the signal that behavior sends about the motives of the agent.

References

- Andersson, P. A., Erlandsson, A., & Västfjäll, D. (2022). Norm avoiders: The effect of optional descriptive norms on charitable donations. *Journal of Behavioral Decision Making*, 35(1), Article e2244. <https://doi.org/10.1002/bdm.2244>
- Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, 100(401), 464–477. <https://doi.org/10.2307/2234133>
- Andreoni, J., & Bernheim, B. D. (2009). Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, 77(5), 1607–1636. <https://doi.org/10.3982/ECTA7384>
- Andreoni, J., Rao, J. M., & Trachtman, H. (2017). Avoiding the ask: A field experiment on altruism, empathy, and charitable giving. *Journal of Political Economy*, 125(3), 625–653. <https://doi.org/10.1086/691703>
- Aquino, K., & Reed, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83(6), 1423–1440. <https://doi.org/10.1037/0022-3514.83.6.1423>
- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology*, 71(2), 364.
- Bandura, A., Caprara, G. V., Barbaranelli, C., Pastorelli, C., & Regalia, C. (2001). Sociocognitive self-regulatory mechanisms governing transgressive behavior. *Journal of Personality and Social Psychology*, 80(1), 125.
- Barclay, P. (2004). Trustworthiness and competitive altruism can also solve the “tragedy of the commons”. *Evolution and Human Behavior*, 25(4), 209–220. <https://doi.org/10.1016/j.evolhumbehav.2004.04.002>
- Bartling, B., Engl, F., & Weber, R. A. (2014). Does willful ignorance deflect punishment? An experimental study. *European Economic Review*, 70, 512–524. <https://doi.org/10.1016/j.eurocorev.2014.06.016>
- Bašić, Z., & Verrina, E. (2021). *Personal norms—And not only social norms—Shape economic behavior (SSRN Scholarly Paper 3720539)*. <https://doi.org/10.2139/ssrn.3720539>
- Batson, C. D., Bolen, M. H., Cross, J. A., & Neuringer-Benefiel, H. E. (1986). Where is the altruism in the altruistic personality? *Journal of Personality and Social Psychology*, 50(1), 212.
- Batson, C. D., & Shaw, L. L. (1991). Evidence for altruism: Toward a pluralism of prosocial motives. *Psychological Inquiry*, 2(2), 107–122.
- Batson, C. D., Thompson, E. R., & Chen, H. (2002). Moral hypocrisy: Addressing some alternatives. *Journal of Personality and Social Psychology*, 83(2), 330.
- Baumeister, R. F. (1998). The self. In (4th ed., Vols. 1–2. *The handbook of social psychology* (pp. 680–740). McGraw-Hill.
- Bazerman, M. H., Tenbrunsel, A. E., & Wade-Benzoni, K. (1998). Negotiating with yourself and losing: Making decisions with competing internal preferences. *Academy of Management Review*, 23(2), 225–241. <https://doi.org/10.5465/amr.1998.533224>
- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), Vol. 6. *Advances in experimental social psychology* (pp. 1–62). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60024-6](https://doi.org/10.1016/S0065-2601(08)60024-6)
- Bénabou, R., & Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review*, 96(5), 1652–1678. <https://doi.org/10.1257/aer.96.5.1652>
- Berman, J. Z., & Small, D. A. (2012). Self-interest without selfishness: The hedonic benefit of imposed self-interest. *Psychological Science*, 23(10), 1193–1199.
- Bicchieri, C. (2005). *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press.
- Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin*, 88, 1–45. <https://doi.org/10.1037/0033-2909.88.1.1>
- Bursztyn, L., & Jensen, R. (2017). Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure. *Annual Review of Economics*, 9(1), 131–153. <https://doi.org/10.1146/annurev-economics-063016-103625>
- Cain, D. M., Dana, J., & Newman, G. E. (2014). Giving versus giving in. *The Academy of Management Annals*, 8(1), 505–533. <https://doi.org/10.1080/19416520.2014.911576>
- Chan, T. Y., Liao, L., Martin, X., & Wang, Z. (2023). Avoiding peer information and its effects on charity crowdfunding: A field experiment. *Management Science*. <https://doi.org/10.1287/mnsc.2023.4807>
- Conrads, J., & Irlenbusch, B. (2013). Strategic ignorance in ultimatum bargaining. *Journal of Economic Behavior & Organization*, 92, 104–115. <https://doi.org/10.1016/j.jebo.2013.05.010>
- Conte, R., Andrighetto, G., & Campenni, M. (2010). Internalizing norms: A cognitive model of (social) norms' internalization. *International Journal of Agent Technologies and Systems*, 2(1), 63–73. <https://doi.org/10.4018/jats.2010120105>
- Dana, J., Cain, D. M., & Dawes, R. M. (2006). What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes*, 100(2), 193–201. <https://doi.org/10.1016/j.obhdp.2005.10.001>
- Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory*, 33(1), 67–80. <https://doi.org/10.1007/s00199-006-0153-z>
- DellaVigna, S., List, J. A., & Malmendier, U. (2012). Testing for altruism and social pressure in charitable giving. *The Quarterly Journal of Economics*, 127(1), 1–56. <https://doi.org/10.1093/qje/qjr050>
- Dunning, D. (2007). Self-image motives and consumer behavior: How sacrosanct self-beliefs sway preferences in the marketplace. *Journal of Consumer Psychology*, 17(4), 237–249. [https://doi.org/10.1016/S1057-7408\(07\)70033-5](https://doi.org/10.1016/S1057-7408(07)70033-5)
- Eckel, C. C., & Grossman, P. J. (1996). Altruism in anonymous dictator games. *Games and Economic Behavior*, 16(2), 181–191.
- Exley, C. L. (2016). Excusing selfishness in charitable giving: The role of risk. *The Review of Economic Studies*, 83(2), 587–628. <https://doi.org/10.1093/restud/rdv051>
- Fehr, E., & Fischbacher, U. (2002). Why social preferences matter – The impact of non-selfish motives on competition, cooperation and incentives. *The Economic Journal*, 112(478), 1–33. <https://doi.org/10.1111/1468-0297.00027>
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3), 817–868. <https://doi.org/10.1162/003355399556151>
- Feiler, L. (2014). Testing models of information avoidance with binary choice dictator games. *Journal of Economic Psychology*, 45, 253–267. <https://doi.org/10.1016/j.joep.2014.10.003>
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.
- Fiske, S. T. (2009). *Social beings: Core motives in social psychology*. John Wiley & Sons.
- Gächter, S., Gerhards, L., & Nosenzo, D. (2017). The importance of peers for compliance with norms of fair sharing. *European Economic Review*, 97, 72–86. <https://doi.org/10.1016/j.eurocorev.2017.06.001>
- Giving USA Foundation. (2022). *Giving USA 2022: The annual report on philanthropy for the year 2021*. <https://givingusa.org/>
- Goffman, E. (1959). The moral career of the mental patient. *Psychiatry*, 22(2), 123–142. <https://doi.org/10.1080/00332747.1959.11023166>
- Grossman, Z. (2014). Strategic ignorance and the robustness of social preferences. *Management Science*, 60(11), 2659–2665. <https://doi.org/10.1287/mnsc.2014.1989>

- Grossman, Z. (2015). Self-signaling and social-signaling in giving. *Journal of Economic Behavior & Organization*, 117, 26–39. <https://doi.org/10.1016/j.jebo.2015.05.008>
- Grossman, Z., & van der Weele, J. J. (2017). Self-image and willful ignorance in social decisions. *Journal of the European Economic Association*, 15(1), 173–217. <https://doi.org/10.1093/jeaa/jvw001>
- Hardy, C. L., & Van Vugt, M. (2006). Nice guys finish first: The competitive altruism hypothesis. *Personality and Social Psychology Bulletin*, 32(10), 1402–1413. <https://doi.org/10.1177/0146167206291006>
- Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, 52(12), 1280.
- Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In *Vol. 2. Advances in experimental social psychology* (pp. 219–266). Elsevier.
- Jordan, J., Leliveld, M. C., & Tenbrunsel, A. E. (2015). The moral self-image scale: Measuring and understanding the malleability of the moral self. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01878>
- Kawamura, Y., Ohtsubo, Y., & Kusumi, T. (2021). Effects of cost and benefit of prosocial behavior on reputation. *Social Psychological and Personality Science*, 12(4), 452–460. <https://doi.org/10.1177/1948550620929163>
- Kimbrough, E. O., & Vostroknutov, A. (2016). Norms make preferences social. *Journal of the European Economic Association*, 14(3), 608–638. <https://doi.org/10.1111/jeaa.12152>
- Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review*, 90(4), 1072–1091. <https://doi.org/10.1257/aer.90.4.1072>
- Kruglanski, A. W. (1989). The psychology of being “right”: The problem of accuracy in social perception and cognition. *Psychological Bulletin*, 106(3), 395.
- Krupka, E. L., & Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, 11(3), 495–524. <https://doi.org/10.1111/jeaa.12006>
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480.
- Lazear, E. P., Malmendier, U., & Weber, R. A. (2012). Sorting in experiments with application to social preferences. *American Economic Journal: Applied Economics*, 4(1), 136–163. <https://doi.org/10.1257/app.4.1.136>
- Lin, S. C., Schaumberg, R. L., & Reich, T. (2016). Sidestepping the rock and the hard place: The private avoidance of prosocial requests. *Journal of Experimental Social Psychology*, 64, 35–40. <https://doi.org/10.1016/j.jesp.2016.01.011>
- Lindsey, L. L. M., Yun, K. A., & Hill, J. B. (2007). Anticipated guilt as motivation to help unknown others: An examination of empathy as a moderator. *Communication Research*, 34(4), 468–480.
- Matthey, A., & Regner, T. (2015). More than outcomes: The role of self-image in other-regarding behavior. *Review of Behavioral Economics*, 2(4), 353–378. <https://doi.org/10.1561/105.000000038>
- Momsen, K., & Ohndorf, M. (2020). When do people exploit moral wiggle room? An experimental analysis of information avoidance in a market setup. *Ecological Economics*, 169, Article 106479. <https://doi.org/10.1016/j.ecolecon.2019.106479>
- Monin, B., & Jordan, A. H. (2009). The dynamic moral self: A social psychological perspective. *Personality, Identity, and Character: Explorations in Moral Psychology*, 341–354.
- Monin, B., & Norton, M. I. (2003). Perceptions of a fluid consensus: Uniqueness bias, false consensus, false polarization, and pluralistic ignorance in a water conservation crisis. *Personality and Social Psychology Bulletin*, 29(5), 559–567.
- Norton, M. I., Vandello, J. A., & Darley, J. M. (2004). Casuistry and social category bias. *Journal of Personality and Social Psychology*, 87(6), 817–831. <https://doi.org/10.1037/0022-3514.87.6.817>
- Ohtsuki, H., & Iwasa, Y. (2006). The leading eight: Social norms that can maintain cooperation by indirect reciprocity. *Journal of Theoretical Biology*, 239(4), 435–444. <https://doi.org/10.1016/j.jtbi.2005.08.008>
- Rachlin, H. (2002). Altruism and selfishness. *Behavioral and Brain Sciences*, 25(2), 239–250.
- Rothmund, T., & Baumert, A. (2014). Shame on me: Implicit assessment of negative moral self-evaluation in shame-proneness. *Social Psychological and Personality Science*, 5(2), 195–202.
- Snyder, M. L., Kleck, R. E., Strenta, A., & Mentzer, S. J. (1979). Avoidance of the handicapped: An attributional ambiguity analysis. *Journal of Personality and Social Psychology*, 37(12), 2297–2306. <https://doi.org/10.1037/0022-3514.37.12.2297>
- Thøgersen, J. (2008). Social norms and cooperation in real-life social dilemmas. *Journal of Economic Psychology*, 29(4), 458–472. <https://doi.org/10.1016/j.joep.2007.12.004>
- Willer, R. (2009). Groups reward individual sacrifice: The status solution to the collective action problem. *American Sociological Review*, 74(1), 23–43. <https://doi.org/10.1177/000312240907400102>
- Woolley, K., & Risen, J. L. (2021). Hiding from the truth: When and how cover enables information avoidance. *Journal of Consumer Research*, 47(5), 675–697. <https://doi.org/10.1093/jcr/ucaa030>
- Xu, Q., He, S., Li, Z., Duan, R., & Li, P. (2023). Voluntary or reluctant? Social influence in charitable giving: an ERP study. *Social Cognitive and Affective Neuroscience*, 18(1), nsad010. <https://doi.org/10.1093/scan/nsad010>