Engaging Charitable Giving: The Motivational Force of Narrative Versus Philosophical Argument

Abstract: Are philosophical arguments as effective as narratives in influencing charitable giving and attitudes toward it? In four experiments, we exposed online research participants to either philosophical arguments in favor of charitable giving, a narrative about a child whose life was improved by charitable donations, both the narrative and the argument, or a control text (a passage from a middle school physics text or a description of charitable organizations). Participants then expressed their attitudes toward charitable giving and were either asked how much they would hypothetically donate if given $10 (Experiment 1) or told they had a 10% chance of winning $10 and given the opportunity to donate from their potential winnings (Experiments 2-4). Across the four experiments, participants in all of the narrative conditions and in some of the argument conditions tended to express more positive attitudes toward charitable giving and donated about $1 more on average than did participants in the control conditions. These effects appear to have been mediated by the "narrative transportation" scale (Green & Bock 2000), which suggests that appeals to donate can be effective if they engage participants' emotions, imagery, and interest.

Key words: charity, empathy, ethics, moral psychology, narrative

Word Count: approx. 12,000 words, plus 9 figures, 2 tables, and online supplements.

Engaging Charitable Giving: The Motivational Force of Narrative Versus Philosophical Argument

1. Introduction

Some of the influences on rates of charitable donation are well-known. For example, people are more likely to donate when specifically asked to do so, when they have reason to think that their donation will be effective, when they are aware that others like them also donate, and when they receive some benefit in return (for a review, see Bekkers & Wiepking 2011). Research participants will also donate more when there is an "identifiable victim" – a specific individual who would be helped by their donation – than when presented with a more general or statistical plea (Lee & Feeley 2016; Kogut & Ritov 2011; Small, Loewenstein, & Slovic 2007). Emotion might also play an important role (including mediating or moderating the identifiable victim effect: Lee & Feeley 2016), though its influence is complicated. Positive mood might enhance giving, but so also might feelings of guilt or sympathy for a victim's situation, and some evidence suggests that positively valenced appeals might have a positive effect on attitudes but be less effective in soliciting actual donations than negatively valenced appeals (Bekkers & Wiepking 2011; Bhati & Hansen 2020; Erlandsson, Nilsson, & Västfjäll 2018; Small & Verrochi 2009).

Charitable organizations don't just present a potential donor with lists of cold facts about the suffering they seek to ameliorate or detailed theories of its underlying causes. They also employ narratives that describe real people who have been helped by the organization's work. Theories and statistics can connect entities and events, but a narrative goes further by telling a story (Velleman 2003). Perhaps "little of substance can be inferred" from the fact that a text

meets the minimal conditions for being a narrative (Lamarque 2004: 394). Nevertheless, canonical narratives—found in novels, films, and short stories—connect people and events together with a thematic unity that engages readers' emotions and attention (Currie 2010: 98). Some philosophers have argued that narratives should have a prominent place in educating participants in a democratic society. Martha Nussbaum (1995) argues that for a democratic society to function well, properly taking into account the lives of its citizens in the creation of institutions, laws, and policies, the people who compose that society must possess a particular sort of skill. This skill, which Nussbaum calls the "literary imagination", grants abilities that, when employed properly, produce a better understanding of the human condition, resulting in people who are better "world citizens".

A substantial body of psychological research suggests that narratives can shift attitudes and behavior. Exposure to sophisticated narratives appears to enhance people's skills at theory of mind and empathetic understanding (Kidd & Castano 2013; Koopman & Hakemulder 2015). Brief stories of moral exemplars can inspire people to donate more to charitable organizations (Han et al. 2022). Narratives have also been used effectively in shifting public opinion and behavior on social issues through telenovelas in Mexico (Singhal & Rogers 1999) and radio soap operas in Rwanda (Paluck 2009). Some research relies on a standard eleven-item "narrative transportation" scale that measures how engaged, involved, or moved one is by a text (Green & Brock 2000; Green & Sestir 2017). For example, studies have found that people who were highly transported by narratives report higher degrees of story-consistent belief changes (Green & Brock 2000), and more positive emotions and evaluative attitudes towards elements depicted in the narrative (Chang 2009; Escalas, Moore, & Britton 2004).

In a different vein, applied ethicists, members of the "effective altruism" movement, and others who emphasize a rational, argument-based approach toward charitable giving often hope that rational arguments might convince people to donate more to charities (MacAskill 2015). However, evidence on this question is limited. As far as we are aware, there are only two published studies that look specifically at whether philosophical argumentation influences charitable giving.

First, Lindauer, Mayorga, Greene, Slovic, Västfjäll, & Singer (2020) found that participants exposed to a 348-word philosophical argument in favor of charitable giving gave away a larger percentage of their possible winnings of $100 than participants who read no text and went straight to a donation prompt. Although this experiment is suggestive, it is unclear how broadly the results would generalize. First, if willingness to donate potential winnings from an experiment depends partly on length and nature of participation in the experiment, results might have looked different if the control condition had also included a reading task requiring similar participant time and engagement. Second, participants might find an unspecified and presumably small "chance of winning $100" less motivating and easier to sacrifice than a set endowment or a specified chance (e.g., a 10% chance of $10). Finally, in preliminary analyses of some of the data reported below (see XXX), we had found null results using stimuli similar to those of Lindauer and collaborators, raising the question of robustness or replicability. Although Lindauer and collaborators' study is preregistered, they report only a single experiment without follow-up and a modest p-value for the main statistical test of their central hypothesis ($p = .036$).

Lindauer and his collaborators have published a second series of studies that overcomes these limitations with mixed results (Buckland, Lindauer, Rodriguez-Arias, and Veliz 2021). In their effort to compare the effects of different types of philosophical arguments on charitable

giving, the researchers included a control text and a clear chance of earning a bonus that could be donated (1/30 chance of $20). In one experiment, they found that participants exposed to two different philosophical arguments for charitable giving donated a bit more to charity than participants in a control condition. However, when they attempted to replicate this result in a second experiment, only one of the two arguments proved effective. More research is warranted, especially direct comparisons of philosophical arguments to narratives.

There are further reasons not to hastily conclude that philosophical arguments are behaviorally effective. Prior research has shown that rational appeals to donate, such as statistical information about poverty, can be somewhat effective on their own but not when mixed with more "emotional" appeals, such as images and descriptions of a particular individual in need (e.g., Small et al. 2007; Erlandsson et al. 2016). Other studies have found that donations were increased by pictures of potential beneficiaries but not information about how much good the donations would generate (Bergh & Reinstein 2021).

Moreover, one might think that if anyone would be influenced by philosophical arguments to give to charity at comparatively high rates it would be professors of ethics who study philosophical arguments and endorse demanding norms of charitable giving. Yet a series of findings from Schwitzgebel's research group suggests that ethics professors do not behave differently, on average, than professors who specialize in other areas of philosophy or professors in departments other than philosophy (Rust & Schwitzgebel 2014; Schwitzgebel & Rust 2016; see also Schönegger & Wagner 2019; Hansson 2020). On charitable giving specifically, ethics professors embrace more demanding norms than do comparison groups, but do not appear to personally give more to charity (Schwitzgebel & Rust 2014). Relatedly, Schwitzgebel and collaborators have found that although philosophical classroom instruction might change

attitudes and behavior concerning the ethics of eating meat, exposure to a philosophical argument for charitable giving had no positive effect and maybe even a "backfire" effect on students' expressed attitudes toward the ethics of giving (Schwitzgebel, Cokelet, & Singer 2020). These studies by the Schwitzgebel group have the advantage of ecological validity but differ substantially in method from those of Lindauer and his collaborators (Lindauer et al. 2020; Buckland et al. 2021), so they are not directly comparable; and there are advantages to tightly controlled online presentations of material.

Questions about the effectiveness, or not, of philosophical argumentation in leading to attitude and behavior change are central to assessing the value of philosophy. Arguably, an aim of philosophical ethical reasoning is not only theoretical understanding, but also, as Aristotle suggests, to actually become good (Aristotle 4th c. BCE/1962). The hope that exposure to philosophical reasoning will tend to improve the future ethical behavior of students is also part of the historical justification for requiring applied ethics classes, especially for business majors (Abend 2014). No simple series of experiments could adequately address complicated questions about the relationship between exposure to philosophical argumentation and real-world moral choice. Nonetheless, empirical evidence ought not to be irrelevant to philosophers' assessments of the effects and potential value of philosophical argumentation. Abundant evidence that philosophical arguments sway ordinary people's attitudes and behavior, for example, would suggest a broader range of potential roles for philosophical reasoning in public, private, and academic contexts than would abundant evidence that philosophical arguments generally have little to no influence on attitudes or behavior.

Although some extant research measures how charitable giving is influenced by a brief description of a person in need versus statistical information about the need, further research is

warranted on the effects of archetypal narratives compared to philosophical arguments. Of course, there is no sharp boundary between the two. Like historical explanations and scientific accounts, philosophical arguments can look more or less like a narrative. The category of narrative resists a precise definition, but we can operate with a graded notion of *narrativity* (Currie 2010: 34; Lamarque 2004). Novels and short stories are high in narrativity while statistics, abstract arguments, and isolated sentences like "Aditi whistled while walking to the market" are all low in narrativity.

In a series of four experiments, we aimed to compare the effects of moving narratives versus philosophical arguments on the charitable attitudes and behavior of online research participants. Our research builds upon existing work on the power of narrative transport as well as extends and attempts to conceptually replicate the work of Lindauer and colleagues. To gauge experienced narrativity, we used the standard measure of narrative transportation developed by Melanie Green and Timothy Brock (Green & Brock 2000). To measure charitable motivation, we developed a five-item charitable motivation scale. In Experiment 1, we measured giving by means of a hypothetical donation. In Experiments 2-4, we offered participants a surprise 10% chance of receiving $10 and presented the opportunity to donate some or all of this possible bonus to one of six effective charities. Experiment 3 also compared the effects of several different narratives and arguments that differed in length and demandingness. Using an improved design, Experiment 4 attempted to determine whether other philosophical arguments could be effective.

2. Experiment 1

2.1 Method

Open Science. Experiment 1 was exploratory and not preregistered, laying the foundation for the preregistration of subsequent experiments. Raw data and stimulus materials for all experiments are available as online supplementary material and at XXXX.

Participants. Mechanical Turk workers (U.S. only, 50+ previous HITs, 95%+ approval rating) were recruited from May 1 to June 8, 2017 to "read a short piece of text and then answer questions about your attitudes and beliefs" for $0.25 compensation, then linked to a survey on Qualtrics. We aimed for a sample size after exclusions of 200 per condition across four conditions. In all, 1220 workers completed the survey, of whom 197 were excluded for failing a comprehension check, leaving 1023 included participants (603 female, 411 male, 9 other or unspecified; median age 30-39), achieving our target sample size in a single recruitment batch.

Design. After consenting, participants were randomly assigned to one of four conditions: control, argument only, narrative only, or argument-plus-narrative. Participants in the control condition read a portion of a middle school physics text on the nature of energy (445 words). Participants in the argument condition read an argument that had been written by professional philosophers specifically with the aim of convincing online research participants to donate to charity (482 words). Participants in the narrative condition read a true account of a child rescued from slavery by a charitable organization (582 words). Participants in the argument-plus-narrative condition read both the argument and the narrative in random order. All participants then answered five questions concerning their attitudes toward charitable giving, were asked hypothetically how much they would donate if given a bonus of $10, and finally answered comprehension and demographic questions.

Stimulus materials. The text for the control condition concerned the nature of energy and was adapted from Hsu (2007).

The text for the argument condition was adapted from Buckland, Lindauer, Rodríguez-Arias, and Véliz (2021) and used with permission. The text was inspired by Peter Singer's arguments for charitable giving, had been written with the aim of convincing research participants to donate to charity, and received Singer's approval for that end. The full text ran as follows:

> The following text argues that we have a moral duty to donate to trustworthy and effective aid agencies that combat poverty. Please read as many times as you need to fully understand the argument. Only click "next" when you feel that you adequately understand the text.
>
> 1. A great deal of extreme poverty exists, which involves suffering and death from hunger, lack of shelter, and lack of medical care. Roughly a third of human deaths (some 50,000 daily) are due to poverty-related causes.
> 2. If you can prevent something bad from happening, without sacrificing anything nearly as important, you ought to do so and it is wrong not to do so.
> 3. By donating money to trustworthy and effective aid agencies that combat poverty, you can help prevent suffering and death from lack of food, shelter, and medical care, without sacrificing anything nearly as important.
> 4. Countries in the world are increasingly interdependent: you can improve the lives of people thousands of miles away with little effort.

5. Your geographical distance from poverty does not lessen your duty to help. Factors like distance and citizenship do not lessen your moral duty.
6. The fact that a great many people are in the same position as you with respect to poverty does not lessen your duty to help. Regardless of whether you are the only person who can help or whether there are millions of people who could help, this does not lessen your moral duty.
7. Therefore, you have a moral duty to donate money to trustworthy and effective aid agencies that combat poverty, and it is morally wrong not to do so.

For example, $20 spent in the United States could buy you a fancy restaurant meal or a concert ticket, or instead it could be donated to a trustworthy and effective aid agency that could use that money to reduce suffering due to extreme poverty. By donating $20 that you might otherwise spend on a fancy restaurant meal or a concert ticket, you could help prevent suffering due to poverty without sacrificing anything equally important. The amount of benefit you would receive from spending $20 in either of those ways is far less than the benefit that others would receive if that same amount of money were donated to a trustworthy and effective aid agency.

Although you cannot see the beneficiaries of your donation and they are not members of your community, it is still easy to help them, simply by donating money that you would otherwise spend on a luxury item. In this way, you could help to reduce the number of people in the world suffering from extreme poverty. You could help reduce suffering and

> death due to hunger, lack of shelter, lack of medical care, and other hardships and risks related to poverty.
>
> With little effort, by donating to a trustworthy and effective aid agency, you can improve the lives of people suffering from extreme poverty. According to the argument above, even though the recipients may be thousands of miles away in a different country, you have a moral duty to help if you can do so without sacrificing anything of equal importance.

While there are many possible arguments for charitable donation with different emphases, this argument in our judgment effectively conveys the central ideas of Singer-style consequentialist arguments of the sort that have been very influential in both academic philosophy and the effective altruism movement. This argument is thus, we hope, representative of the most prominent academic philosophical argumentative approach toward motivating charitable giving.

The text for the narrative condition was the true story of Mamtha copied verbatim from the International Justice Mission charitable website. We thought it desirable to use a true story from a real charitable organization website, both for realism and because it seems likely that charitable organizations' websites reflect those organizations' best practical assessments of what materials successfully drive donations. After reviewing about a hundred publicly available narratives, we chose the Mamtha narrative because its length approximately matched that of the argument and the experimenters and pilot participants judged it to be engaging, well-written, emotionally effective, and broadly typical of the inspirational narratives used by charitable organizations. Mamtha is a ten-year-old girl whose family is tricked into moving from their farm

to a sand mine where they are kept as slaves under harsh conditions and forced to work sixteen hour days. A charitable donation funded a rescue operation that shut down the sand mine, freeing her and her family. Mamtha is now studying to be a doctor. (The full text is available in the supplementary online materials.)

In highlighting a single individual, the Mamtha narrative resembles the type of stimulus commonly used in studies of the identifiable victim effect. However, we note two important differences. First, the Mamtha narrative is much longer and more detailed than the stimuli typically used in identifiable-victim studies. Second, these studies typically promise that the participant's donation will help the particular individual identified, whereas it is clear from our text that Mamtha has already been helped. Therefore, any donation our participants make later in the experiment will not be going to her in particular.

Attitude and donation questions. After indicating that they understood the target text, participants answered the following five attitude questions on a seven-point scale from "I completely disagree" (-3) to "I completely agree" (+3), with the midpoint marked "I neither agree nor disagree".

(1.) It is morally good to give money to charities that help those in extreme poverty.

(2.) People like me should give money to charities that help people in extreme poverty.

(3.) Currently I feel motivated to give money to a charity that helps people in extreme poverty.

(4.) Right now I have no desire to give money to a charity that helps people in extreme poverty. [Reverse scored]

(5.) I think it is important to support charities that help those struggling with extreme poverty.

As a final measure of charitable motivation, and laying the groundwork for Experiments 2-4, we asked participants what they would do with a hypothetical donation:

> Hypothetically, suppose we gave you an additional $10 for participation in this study, along with the option to donate some portion of it to one of six well-known charities that have been shown to effectively fight suffering due to extreme poverty. In this hypothetical case, how much of your additional $10 do you think you would donate?

Response options ran from $0 to $10 in $1 increments.

Narrative transport measure. Typically, texts high in narrativity strongly engage readers' emotions and attention (Velleman 2003; Currie 2010). We operationalize experienced narrativity in our participants by using the eleven-question "narrative transportation" measure from Green and Brock (2000), with some minor modifications (e.g., replacing occurrences of "the story" with "the text", since the control and argument texts are not stories). Sample questions include:

- "While I was reading the text, I could easily picture the events in it taking place,"
- "I wanted to learn how the text ended,"
- "The text affected me emotionally," and
- "While I was reading the text, activity going on in the room around me was on my mind" (reverse scored).

Each item used the same 7-point response scale as for the attitude questions.

2.2 Results

Attitude scale. The five questions constituting the attitude scale showed high intercorrelation, with a Cronbach's alpha of .87. Principal Component Analysis found that a single factor

explained 69% of the variance with each factor loading at least .77. For subsequent analysis, we averaged the variables to create a single attitude score ranging from -3 to +3. As expected, expressed attitude correlated well with hypothetical donation: $r = .43$ ($p < .001$).

Attitude and hypothetical donation by condition. Figures 1 and 2 show differences by condition in both attitude and hypothetical donation. Participants in the narrative (N) and narrative-plus-argument (NA) conditions expressed significantly more positive attitudes toward charitable giving than did participants in the control (C) condition and the argument (A) condition; and argument alone was not statistically different from control, with overall mean attitude scores of 1.77, 1.70, 1.38, and 1.35 respectively (pooled $SD = 1.20$, $F[3, 1019] = 8.33$, $p < .001$; Tukey post-hoc N > A, N > C, NA > A, NA > C).

The results for hypothetical donation showed a very similar pattern with means of $6.21, $6.38, $4.86, and $5.33, respectively (pooled $SD = 3.22$, $F[3, 1019 = 12.4$, $p < .001$; Tukey post-hoc N > A, N > C, NA > A, NA > C). Responses somewhat violated normality, clustering at $5 and $10 (318 and 247 out of 1023). We confirmed the main results non-parametrically with a Mann-Whitney comparison of the medians of the narrative conditions versus the argument and control conditions (median N & NA = 6, median A & C = 5, U = 103929, p < .001).

Figure 1: Mean attitude toward charitable giving by condition, on a scale from -3 to +3, Experiment 1. Error bars represent 95% confidence intervals.

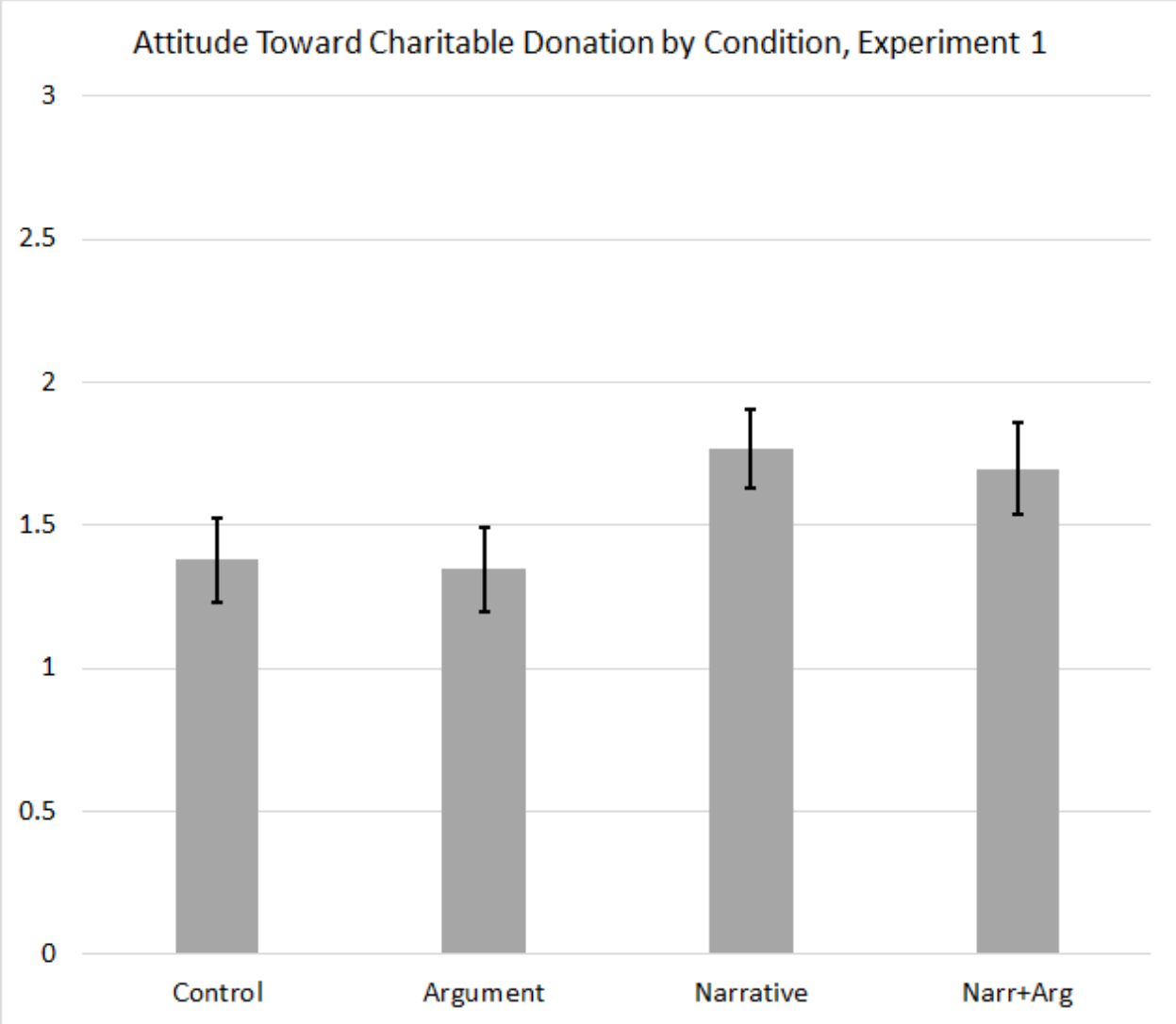
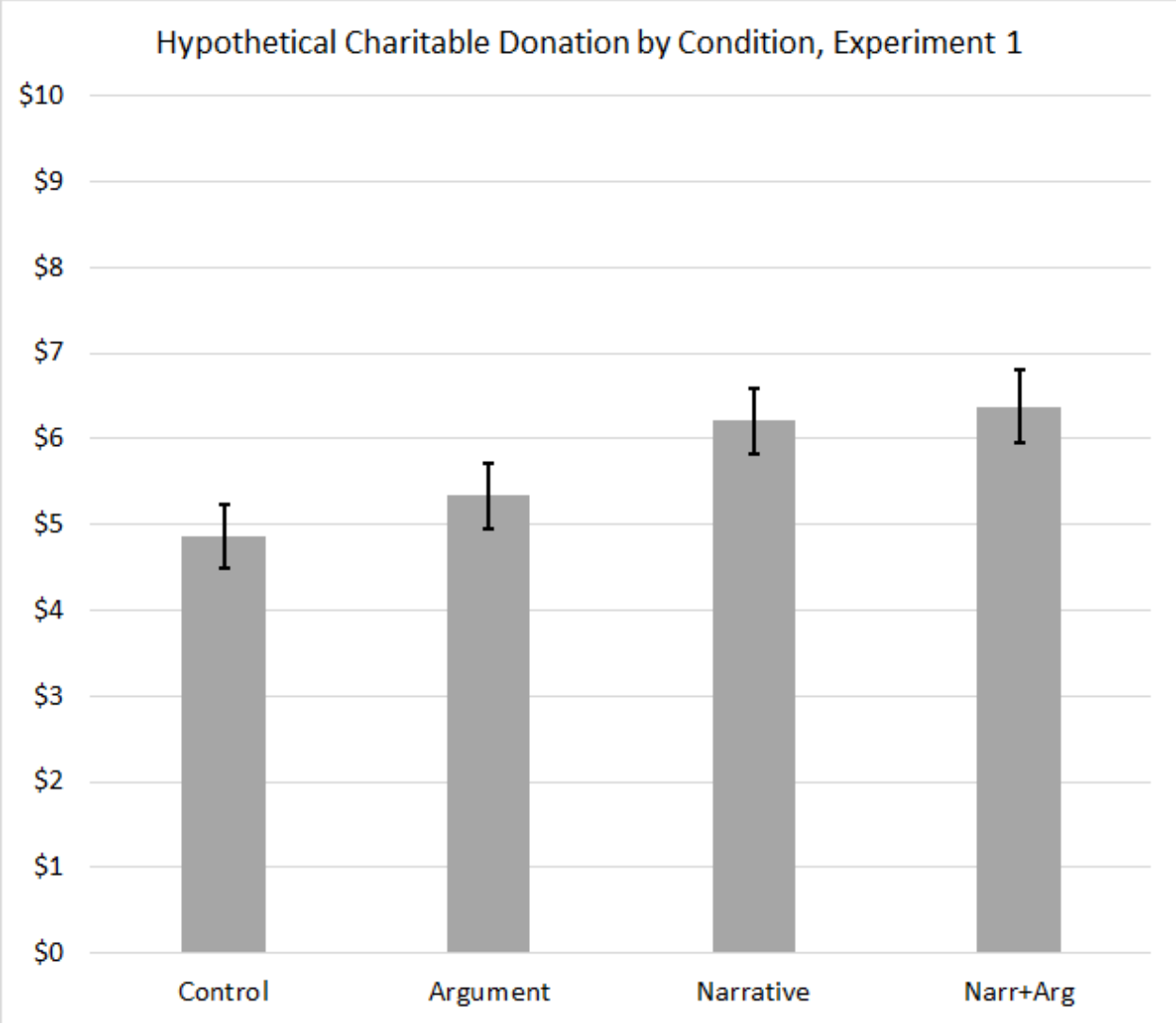

Attitude Toward Charitable Donation by Condition, Experiment 1
3
2.5
2
1.5
1
0.5
0
Control
Argument
Narrative
Narr+Arg

Figure 2: Mean hypothetical charitable donation by condition, Experiment 1. Error bars represent 95% confidence intervals.



Narrative transport. Cronbach's alpha for the 11 transport questions was acceptable at .77, so responses were averaged for an overall transport score from -3 to +3. As expected, mean narrative transport was highest in the narrative conditions, and it was higher in the argument condition than in the control condition: M_N = 1.16, M_{NA} = 1.11, M_A = 0.73, M_C = 0.48 (pooled *SD* = 0.85, $F[3, 1019] = 35.4$, $p < .001$; Tukey post-hoc all differences significant except N vs. NA). Also as expected, transport correlated well with hypothetical donation in all three of the narrative or argument conditions (*r*'s from .32 to .35, *p*'s < .001) but not in the control condition ($r = .10$, $p = .13$). Transport correlated even more strongly with expressed attitude (*r*'s from .58 to .65, $p < .001$). Transport and expressed attitude were also correlated in the control condition, though to a lesser extent, possibly reflecting participant engagement ($r = .27$, $p < .001$).

Mediation analysis. Details of mediation and regression analyses are in Appendix A. In Experiment 1, we found that narrative transport mediated both attitude and hypothetical donation, while a direct effect of assignment to the narrative condition was only detectable for hypothetical donation. This is consistent with a causal model according to which exposure to narrative causes narrative transport which in turn causes more positive attitudes toward charitable donation and higher rates of hypothetical giving. However, other causal models are also possible.

2.3 Discussion

As expected, participants exposed to the narrative expressed more positive attitudes toward donation than did participants in the other conditions, and in a hypothetical donation they said they would give more. Also as expected, the effects of exposure to the narrative on attitude and

hypothetical donation appear to be at least partly mediated by narrative transport. Given prior research on charitable giving (e.g., Lindauer et al. 2020), we were somewhat surprised to see no detectable difference between the argument and control conditions. It is also somewhat surprising that there was no significant difference between the narrative and narrative-plus-argument conditions, in light of prior studies on charitable giving that have reported interesting results from mixing conditions (e.g., Small et al. 2007; Erlandsson et al. 2016; Bergh & Reinstein 2021). However, this initial study did not include a measure of actual donation behavior.

Our primary concern is with the effect of narratives versus arguments on real charitable giving, rather than intentions to donate in a hypothetical scenario. So we conducted three additional studies using a lottery design that more closely resembles charitable giving.

3. Experiment 2

3.1 Method

Open Science. The main methods, hypotheses, and analyses were preregistered at the Open Science Foundation: XXXXX.

Participants. From Jun. 7-21, 2018, we recruited participants through Mechanical Turk with the same methods and criteria as in Experiment 1. However, due to a combination of chance variation and higher exclusion rates in the narrative-plus-argument condition, that condition did not meet the participation threshold. Thus, from Aug. 2-14 we added 54 new narrative-plus-argument participants in a follow-up recruitment to achieve the target sample for that condition. In all, 917 participants completed the study after exclusions (530 female, 386 male, 1 other or unspecified; median age 30-39).

Design. The design was identical to Experiment 1, except with a lottery chance of a real donation instead of a hypothetical donation, an option to choose a charity to receive the donation, and an option to provide an email address to a charitable organization for follow-up contact.

Stimulus materials. The control, argument, and narrative texts were identical to those used in Experiment 1, as were the narrative transport questions. The five attitude questions were identical, except that we changed the wording from "extreme poverty" to "poverty", to better reflect the mission of the range of charitable organizations among which participants would select.

> The hypothetical donation question was replaced by the following *lottery donation*:
>
> Upon completion of this study, 10% of participants will receive an additional $10. You have the option to donate some portion of this $10 to one of six well-known charities that have been shown to effectively fight for suffering due to poverty. If you are one of the recipients of the additional $10, the portion you decide to keep will appear as a bonus credited to your Mechanical Turk account, and the portion you decide to donate will be given to the charity you pick from the list below. If you are one of the recipients of the additional $10, how much of your additional $10 would you like to donate?

Participants were again asked to indicate on an 11-point scale at $1 intervals from $0 to $10 how much they pledge to donate if selected as a winner of the bonus. Participants were then asked "Which charity would you like your chosen donation to go to?" and given the option of Feeding America, National Federation of the Blind, Against Malaria Foundation, Global Alliance for Improved Nutrition, Homes for Our Troops, or Helen Keller International, with a link to each organization's websites. Below the choice, we provided brief charity mission

statements drawn from the organizations' promotional materials. These six organizations were selected in part because of their effectiveness and in part to offer a mix of U.S. versus international charities. Participants were also invited to explain their choice in a free response box.

Finally, participants were invited to optionally share their email with our funding organization, The Life You Can Save, "to receive information on how you might help to effectively fight extreme poverty around the world". For privacy, these responses were unlinked from the questionnaire responses, but we tracked the number of properly formatted email addresses by condition as a final dependent measure of charitable motivation.

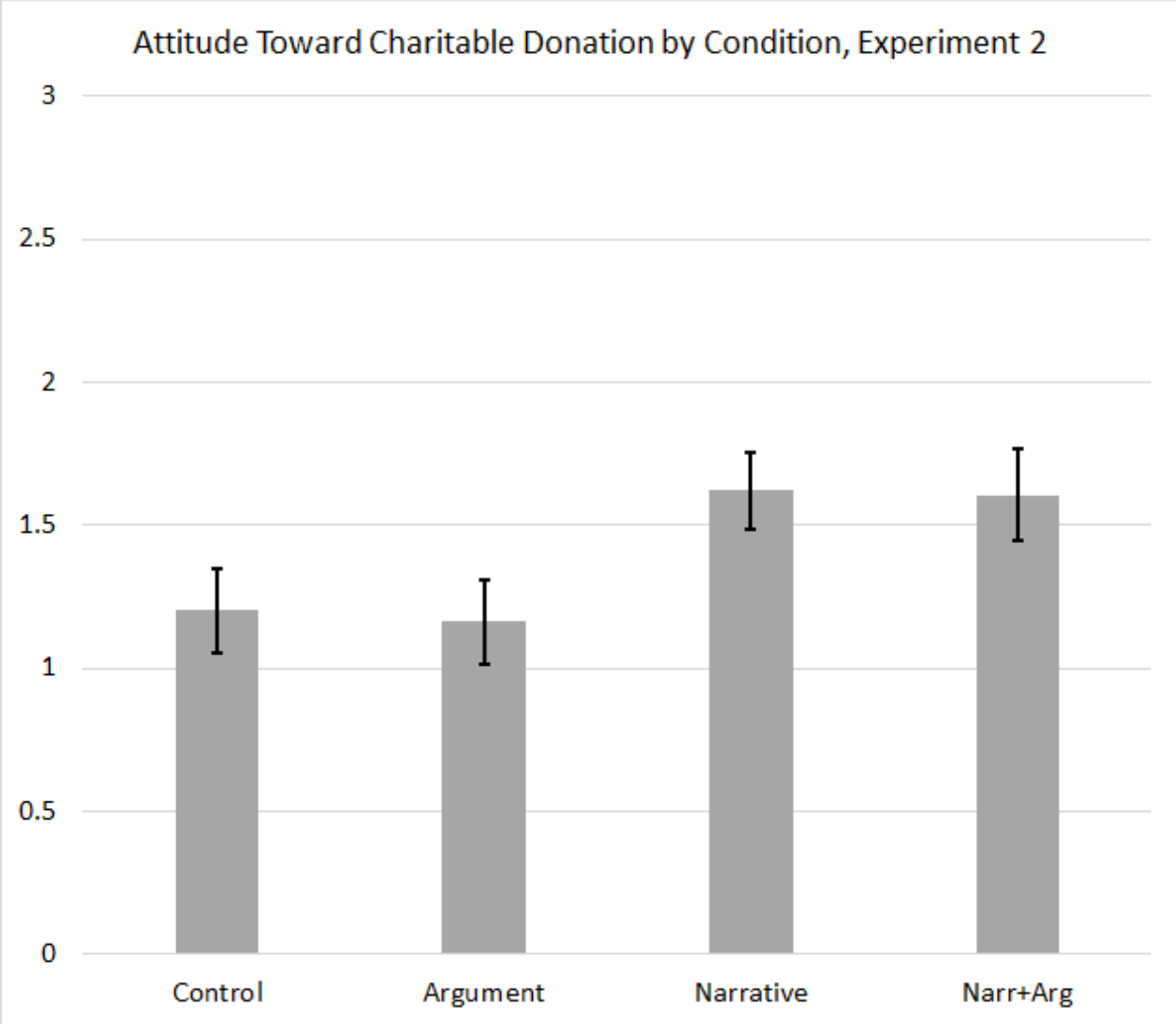
3.2 Results

Attitude scale. The five questions constituting the attitude scale again showed good intercorrelation, with a Cronbach's alpha of .79. Principal Component Analysis found that a single factor explained 61% of the variance, with the reverse-scored question loading at .47 and the other questions loading at .82-.86, so responses were again summed for a single attitude score. Encouragingly, expressed attitude predicted the lottery donation almost as well as it predicted hypothetical donation in Experiment 1: $r = .40$ ($p < .001$).

Attitude by condition. Figure 3 shows expressed attitude by condition. The results replicated the results of Experiment 1. Participants in the narrative (N) and narrative-plus-argument (NA) conditions expressed significantly more positive attitudes toward charitable giving than did participants in the argument (A) condition and the control (C) condition; and argument alone was not statistically different from control, with overall mean attitude scores of 1.62, 1.60, 1.16, and 1.20 respectively (pooled $SD = 1.16$, $F[3, 913] = 10.8$, $p < .001$; Tukey

post-hoc N > A, N > C, NA > A, NA > C). We detected no difference in mean attitude between participants in the main batch NA condition and the participants added to NA in the supplementary recruitment (M_{main} = 1.65 [N = 160] vs. M_{supp} = 1.48 [N = 54], pooled SD = 1.14, t = 0.95, p = .34).

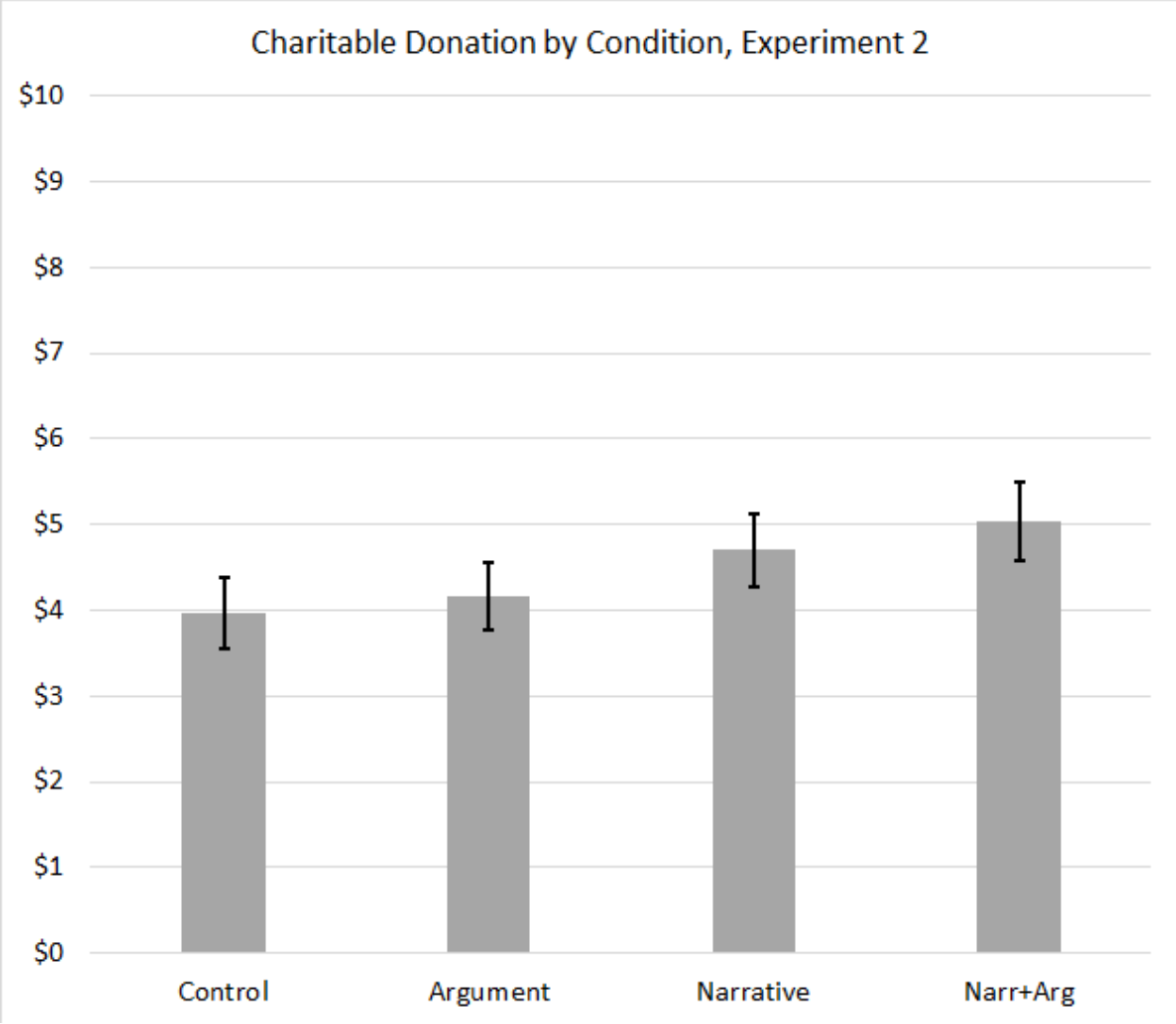
Figure 3: Mean attitude toward charitable giving by condition, on a scale from -3 to +3, Experiment 2. Error bars represent 95% confidence intervals.



Donation by condition. Figure 4 shows donation amount by condition. The overall donation pattern was similar to that in Experiment 1 (though about a dollar lower), with means of $4.70 for narrative, $5.04 for narrative plus argument, $4.17 for argument, and $3.97 for control (pooled *SD* = 3.24, $F[3, 913] = 4.99$, $p = .002$; Tukey post-hoc NA > A, NA > C). In this analysis, the narrative condition alone was not statistically significantly different from either argument alone or the control condition. However, pre-registered one-tailed pairwise t-tests comparing Narrative versus Argument and Control did show statistical significance (M_N = $4.70,

M_A = \$4.17, pooled SD = 3.27, t = 1.78, p = .038; M_N = \$4.70, M_C = \$3.97, pooled SD = 3.13, t = 2.42, p = .008). As in Experiment 1, responses moderately violated normality with clustering at \$0, \$5, and \$10 (144, 272, and 120 out of 917). A Mann-Whitney test of the two narrative conditions (N and NA) versus the non-narrative conditions (A and C) showed identical medians of 5, but nonetheless a statistically significant difference in mean ranks (U = 90610, mean rank = 492 [narrative] vs. 430 [non-narrative], $p < .001$).

Figure 4: Mean donation of a 10% chance of $10 by condition, Experiment 2. Error bars represent 95% confidence intervals.



Email response by condition. Overall, only 9.6% of participants (88/917) offered what appeared to be a valid email address, and no difference by condition was statistically detectable ($\chi^2[3] = 1.8$, p = .61).

Narrative transport. Cronbach's alpha for the 11 transport questions was acceptable at .73, so responses were again summed for an overall transport score from -3 to +3. As expected, narrative transport was highest in the narrative conditions, and it was higher in the argument condition than in the control condition: $M_N = 1.03$, $M_{NA} = 0.92$, $M_A = 0.65$, $M_C = 0.33$ (pooled $SD = .89$, $F[3, 913] = 26.4$, $p < .001$; Tukey post-hoc all differences significant except N vs. NA). As in Experiment 1, transport correlated more strongly with attitude in the experimental conditions (r's from .46 to .57, $p < .001$) than in the control condition ($r = .37$, $p < .001$). However, unlike Experiment 1, transport correlated moderately with donation in all four conditions (r's from .18 to .28, p's $\leq .002$), rather than correlating with all but less in the control condition.

Mediation analysis. See Appendix A. As in Experiment 1, narrative transport mediated both attitude and donation. A direct effect of assignment to the narrative condition was detectable only for attitude.

3.3 Discussion

Experiment 2 replicated the main results from Experiment 1, but now with real stakes of a 10% chance of $10, instead of a hypothetical donation. Participants in the narrative conditions

expressed more positive attitudes toward donation and donated more money than did participants in the control condition or exposed to the philosophical argument alone.

Although our narrative and argument were each chosen to be good exemplars of their type, we wondered whether something about the particular narrative or particular argument might be driving the effect. So, in the next experiment, we expanded the study to include four different narratives and four different arguments. Charitable organizations often argue only that it's morally good to donate, not that it's morally bad not to donate or that you have a duty to donate, which might be experienced as threatening or demanding. Also the formal structure of the argument with the numbered premises might have been off-putting, so we wanted to rewrite the arguments in something closer to ordinary prose.

Since the narratives on charitable websites typically also include pictures of the people helped, we also wanted to see if inclusion of these pictures would amplify the effect. Under some conditions, researchers have found that adding a photograph to a charitable appeal can increase rates of giving (e.g., Genevsky, Västfjäll, Slovic, & Knutson 2013; though see Hart, Lane, & Chinn 2018).

4. Experiment 3

4.1 Method

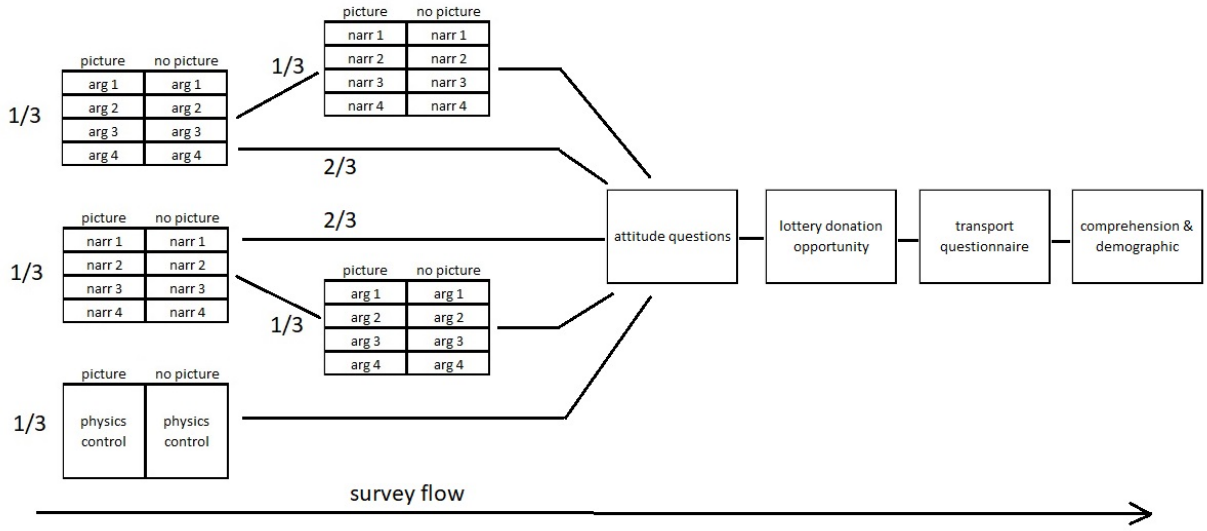
Open Science. The main methods, hypotheses, and analyses were preregistered at AsPredicted.org (#XXXX). The stimulus materials and raw data for all experiments are available in the online supplement and at XXXX.

Participants. We recruited participants through Mechanical Turk (restricted to participants in the U.S. with 100+ HITs and approval ratings of 98%+), from Apr. 11-15, 2019.

In the first phase, 1004 participants completed the study. Due to a survey flow error, the narrative-plus-argument condition was underfilled, so 60 additional participants were added to that condition. Of the 1064 participants who completed the survey, 111 were excluded for failing the comprehension checks or completing in under 2.5 minutes, leaving 953 included participants (513 female, 431 male, 9 other or unspecified; median age 37).

Design. One third of participants were randomly shown either a narrative (one of four), an argument (one of four), or the physics text used in the previous experiments. Among those shown one of the four narratives, one third were then shown one of the four arguments. Among those shown one of the four arguments, one third were then shown one of the four narratives. All participants then saw the five attitude questions, the lottery donation opportunity, the transport questionnaire, the comprehension checks, and then finally demographic information and the opportunity to share their email address with The Life You Can Save. Half of participants also viewed a picture. In the narrative condition, it was a real photo of the protagonist of the narrative. In the argument condition, it was one of the four photos from the narrative stimuli. In the control condition, it was the physicist James Prescott Joule. The overall design of the experiment is shown in Figure 5.

Figure 5: Design of Experiment 3



Stimulus materials. The four narratives ranged in length from 503-782 words and were true stories drawn from the websites of charitable organizations. All stories centered on a protagonist who was a child or teen, described how their lives had been improved, and ended with a hopeful or inspiring tone. Participants in the photo conditions also viewed a color photo of the protagonist, drawn from the same websites. One argument was the same Singer-style argument used in Experiments 1 and 2, which was relatively short, included numbered premises, drew primarily on utilitarian principles, and concluded that participants have a duty to donate. The other three varied either in their *injunction* (concluding that it's good to donate rather than that you have a duty to donate) or *breadth* (in being longer and including reasons for donating that were both utilitarian and non-utilitarian; see Table 1). Participants in the argument photo condition randomly saw one of the four photos from the narrative condition alongside the argument, with the caption "a real person who has been aided by charitable donations".

Table 1: Excerpts from the four types of argument

	Injunction (Good vs. Duty)	
Breadth (Short Utilitarian Only vs. Long Non-Utilitarian Too)	"...you have a **moral duty** to help distant strangers who are suffering due to poverty, by donating to an effective charity, if you can do so without sacrificing anything of equal importance."	"...it is **morally good** to help distant strangers who are suffering due to poverty, by donating to an effective charity, if you can do so without sacrificing anything of equal importance."
	PLUS "You have participated in the system that has produced and maintained poverty and you have benefited from it. … [so] it is your **duty** to help and it is morally wrong if you do not help."	PLUS "You have participated in the system that has produced and maintained poverty and you have benefited from it. … [so] it is **good** for you to help if you have the opportunity to do so."

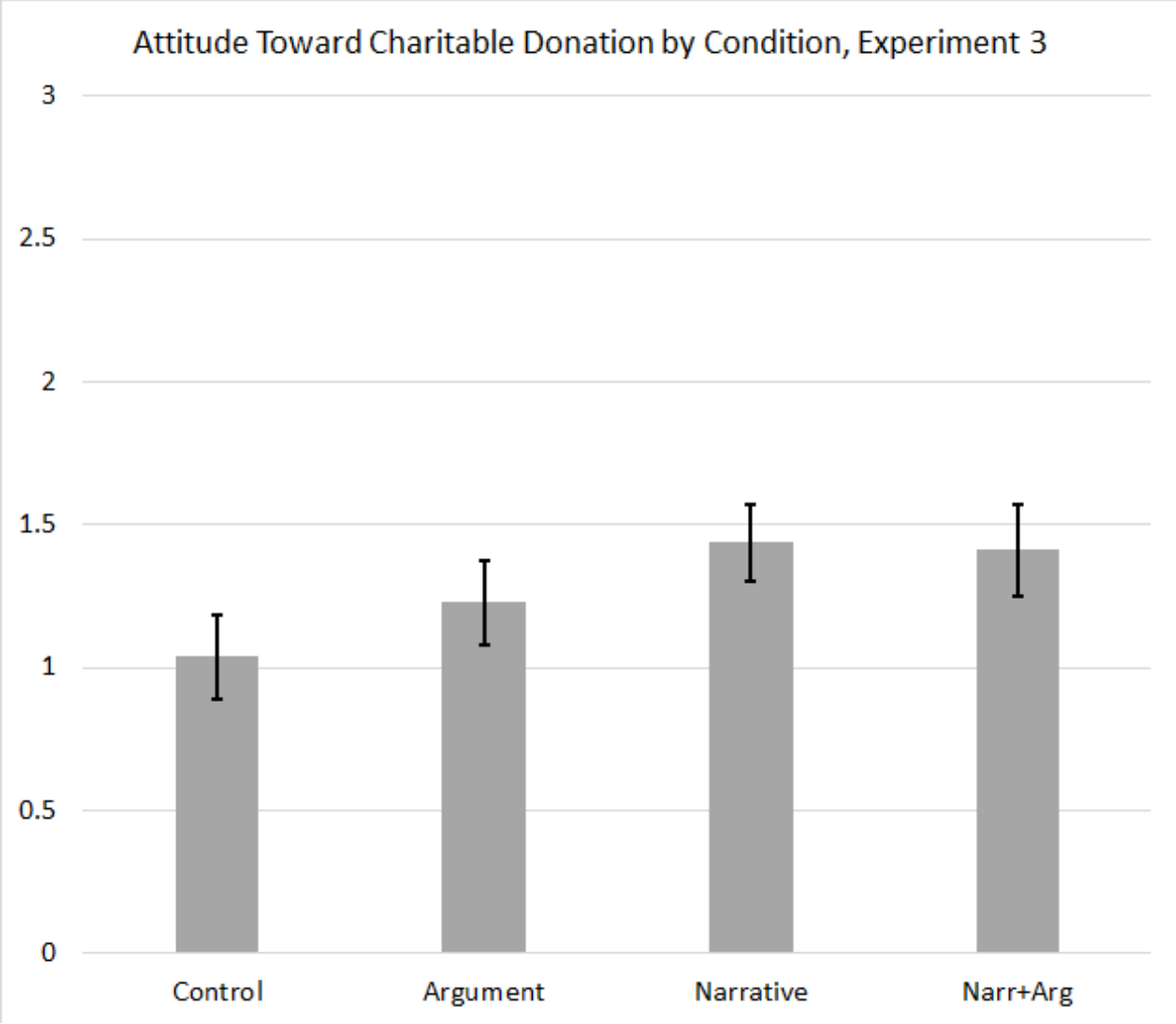
The attitude and transport questions were identical to those in Experiment 1. To address the possibility that some participants might think their odds of receiving the $10 might be increased if they chose to donate, and to encourage careful responding, we added the following text to the lottery donation question: "Note: You must pass the comprehension questions and show no signs of suspicious responding to receive the $10. Receipt of the $10 is NOT conditional, however, on your attitudes toward charity, expressed on the previous page, nor on how much you choose to donate if you receive the $10."

4.2 Results

Attitude scale. The five questions constituting the attitude scale once again showed good intercorrelation, with a Cronbach's alpha of .88, a single factor explaining 70% of the variance in a Principal Component Analysis, and a minimum factor loading of .79. Replicating Experiments 1 and 2, expressed attitude again correlated well with donation: $r = .44$ ($p < .001$).

Attitude by condition. Figure 6 shows attitude by condition. As in Experiments 1 and 2, participants in the narrative conditions expressed more positive attitudes toward charitable giving than in the control condition but the argument did not differ from control. Overall mean attitude scores were 1.44 in the narrative condition, 1.41 in narrative-plus-argument, 1.23 in argument, and 1.04 in control (pooled SD = 1.22, F[3, 949] = 5.69, p = .001; Tukey post-hoc N > C, NA > C). A preregistered t-test found that attitude was significantly more positive in the narrative conditions than in the argument condition (M_{NA+N} = 1.42, M_A = 1.23, pooled SD = 1.25, t(661) = 1.99, p = .024). We detected no difference in mean attitude between participants in the main batch narrative-plus-argument condition and the participants added to this condition in the supplementary recruitment (M_{main} = 1.41 [N = 147] vs. M_{supp} = 1.39 [N = 46], pooled SD =1.31, t(191) = 0.10, p = .92).

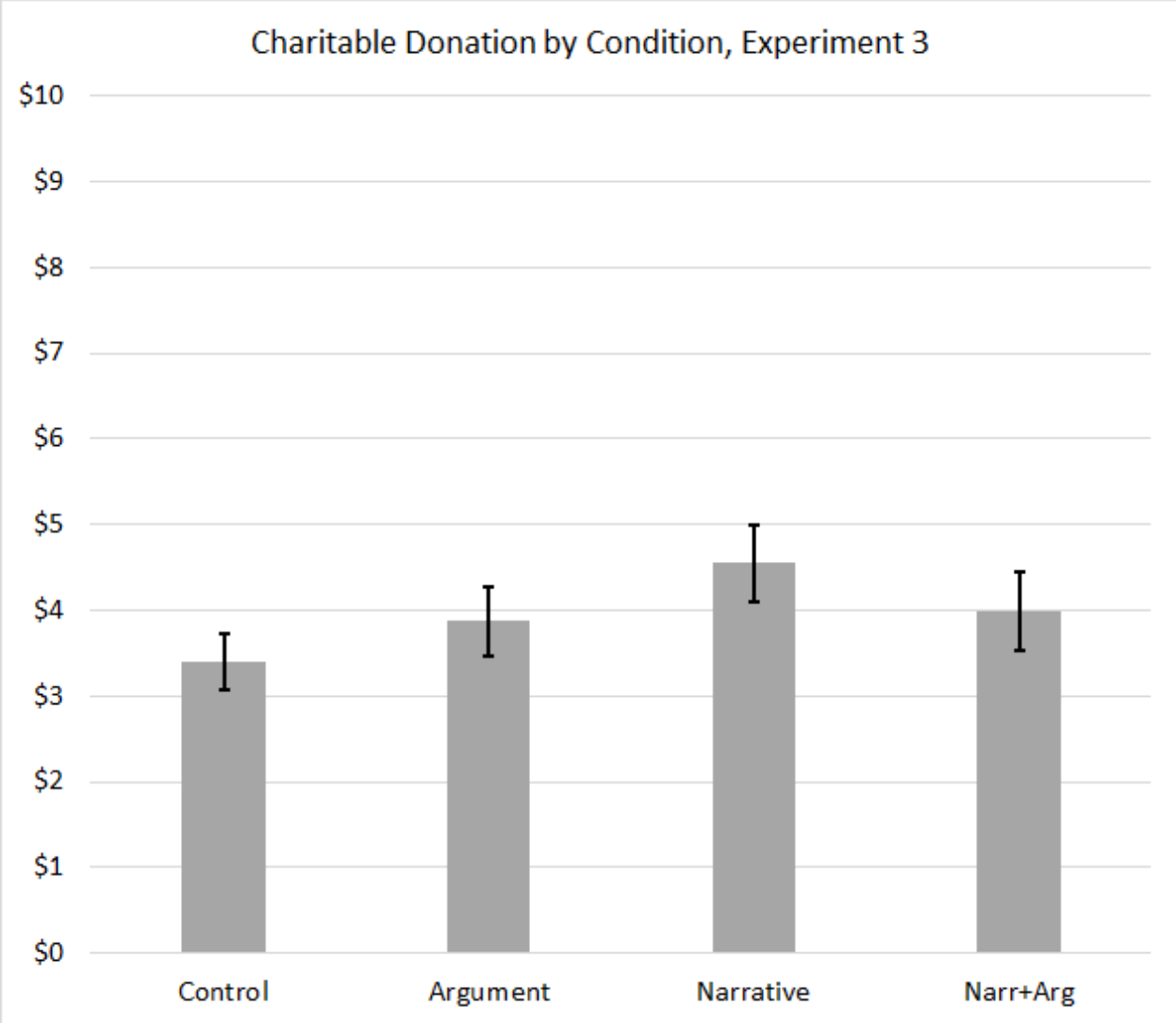
Figure 6: Mean attitude toward charitable giving by condition, on a scale from -3 to +3, Experiment 3. Error bars represent 95% confidence intervals.



Donation, email response, and transport by condition. Figure 7 shows donation amount by condition. The overall donation pattern was similar to that in Experiments 1 and 2, with means of $4.55 for narrative, $4.02 for narrative plus argument, $3.88 for argument, and $3.39 for control (pooled SD = 3.19, F[3, 949] = 5.47, p = .001; Tukey post-hoc N > C only). In this part of the experiment, narrative did not statistically outperform argument, although the narrative condition was the only condition that statistically outperformed control. Clustering was similar to that in the previous experiments.

Email responses were again low at 10.9% (104/953) and did not detectably differ by condition. The narrative transport results were also essentially the same as in Experiments 1 and 2 (Cronbach's alpha = .76; M_N = 0.90, M_{NA} = 0.83, M_A = 0.59, M_C = 0.26 (pooled SD = 0.89, F[3, 949] = 26.5, p < .001; Tukey post-hoc all differences significant except N vs. NA). As in Experiment 1, transport correlated better with donation in the argument and narrative conditions (r = .26 to .39, p's < .001) than in the control condition (r = .14, p = .017).

Figure 7: Mean donation of a 10% chance of $10 by condition, Experiment 3. Error bars represent 95% confidence intervals.



Picture. Including a picture had no detectable effect on the results. Overall, the mean attitude score was very similar whether or not a picture was included in the stimulus (M_{pic} = 6.26, M_{nopic} = 6.14, pooled SD = 6.14, $t(951) = 0.04$, $p = .97$). Excluding the control condition does not alter the lack of effect, nor did any one of the four beneficiary pictures detectably differ from any of the others in pairwise comparisons ($|t| \leq 1.61$, $p \geq .11$). Donation amount was similarly unrelated to the presence or absence of a picture. We had sufficient statistical power for an 80%

chance of detecting a small effect size of Cohen's d = .18 in either mean attitude or mean donation, had a difference been present in the population.

Comparing the four narratives. We had limited statistical power to test for differences among the four narratives. We varied the narratives not to test them individually but to confirm that the results obtained with the Mamtha narrative in Experiments 1 and 2 would generalize to other similar narratives with different protagonists and content. Indeed, the particular narrative did not appear to matter to the results. Mean expressed attitude did not detectably differ (M = 1.36 to 1.54, pooled SD = 1.25, $F(3, 395) = 0.4$, p = .74) nor did mean donation (M = \$4.11 to \$4.41, pooled SD = 3.29, $F(3, 395) = 0.2$, p = .93). We had sufficient statistical power for an 80% chance of detecting a small-to-medium effect size of f = .17 in either mean attitude or mean donation, had a difference been present in the population. Nor did the narratives detectably differ in overall transport (pooled SD = 0.92, $F(3, 395) = 0.5$, p = .66).

Comparing the four arguments. With a 2x2 design, we had the power to detect medium-sized differences between arguments emphasizing *duty* versus arguments concluding only that donation is *good* and also between *shorter* utilitarian arguments and *longer* arguments that added non-utilitarian considerations. As with the narratives, however, our primary aim was not to test for differences among arguments but to confirm that the results of Experiments 1 and 2 would replicate with a somewhat broader range of arguments rather than depending too sensitively on the details of a single argument. The length of the argument had no detectable effect on either attitude or donation (attitude: M_{short} = 1.40, M_{long} = 1.21, pooled SD = 1.27, $t(455) = 1.61$, p = .11; donation M_{short} = \$3.69, M_{long} = \$4.16, pooled SD = 3.31, $t(455) = -1.52$, p = .13). Participants who read the duty argument might have expressed less positive attitudes toward charitable donation, but we interpret the result cautiously given multiple comparisons and a

modest *p* value (attitude: M_{good} = 1.44, M_{duty} = 1.18, pooled *SD* = 1.27, *t*(455) = 2.20, *p* = .029; donation M_{good} = 3.81, M_{duty} = 4.06, pooled *SD* = 3.32, *t*(455) = -0.83, *p* = .41).

Mediation and regression analysis of transport. We created mediation and regression analyses parallel to those in Experiments 1 and 2, ignoring photo, argument type, and narrative protagonist (see Appendix A). As in Experiments 1 and 2, narrative transport mediated both attitude and donation. There was no statistically detectable direct effect of assignment to the narrative condition.

4.3 Discussion

Experiment 3 confirmed the results of Experiments 1 and 2 using a broader selection of narratives and arguments. Supplementing the narrative or argument with a picture had no detectable effect on expressed attitude or donation amounts. We found tentative evidence that phrasing the argument's conclusion as donating is good (as opposed to a duty) had a positive effect on attitude toward charity, but this was not supported by any detectable difference in donation amounts.

Together, our experiments so far have suggested that standard narratives are more effective than standard philosophical arguments at promoting more charitable giving. However, the experimental design of the previous studies could be improved. We failed to include a clear *manipulation check*, to make sure participants perceived the narratives as narratives and arguments as arguments. We also always presented the *attitude questions first*, which might influence donation pledges, through either an order or demand effect. The longer stimulus in the narrative-plus-argument conditions required more participant effort, possibly explaining the higher exclusion rates in this condition than in the other conditions in Experiment 2. The added

length of this stimulus might also have been a confounding factor by affecting either participant investment in the experiment (which might increase donation rates relative to the other conditions) or participants' sense that they deserve higher payment (which might decrease donation rates relative to the other conditions). Finally, our control and manipulation conditions differ not only in terms of having a narrative or argumentative structure but also in making charitable donation salient (similar to the "non-moral argument" control condition in Buckland et al. 2021). It would be ideal to have a *new control condition* in which charitable organizations are described, just not in terms of an argument in favor of donating to them or a narrative in which individuals in need are helped.

The previous studies also leave open two important questions. First, are philosophical arguments *always* less moving than narratives, or can they be written to perform as well (or better)? Other recent research suggests that philosophical arguments can be effective at increasing charitable donations (e.g., Lindauer et al. 2020; Schwitzgebel & Cushman 2020; Buckland et al. 2021). Indeed, philosophical arguments are often paired with thought experiments that take on something of a narrative form—such as Peter Singer's famous example in which a passerby merely has to ruin his shoes in order to save a young child drowning in a pond. Second, when narratives are more effective than standard philosophical arguments, *why* is that so? Our mediation analyses suggest that narratives promote greater narrative transport. Perhaps more persuasive arguments are more effective because they have higher narrativity. In our final experiment, we test this hypothesis and improve our experimental design and materials.

Between Experiment 3 and Experiment 4, we ran another pre-registered experiment (Experiment 3a) aimed at varying the level of "threat" in the arguments (by contrasting arguments that "you" should donate versus that "the wealthy" should donate). In retrospect, we

believe the threat manipulation failed, so the experiment is not included here in the main text. However, since the experiment replicated the main results of Experiments 1-3 and some readers might find it of interest, we describe its methods and results in Appendix B.

5. Experiment 4

5.1 Method

Open Science. We preregistered the main methods, hypotheses, and analyses at XXXX.

Participants. We recruited participants through Prolific (restricted to users over the age of 19 residing in the USA or UK with a minimum approval rate of 95+%) to complete a survey on Qualtrics for $1.50, on Mar. 8-9, 2022. Of the 1500 participants who completed the survey, 68 were excluded for failing comprehension checks or completing in under 2.5 minutes, leaving 1432 included participants (879 female, 533 male, 20 other or unspecified; median age 38).

Design and Stimulus Materials. Participants were randomly assigned to one of five conditions: *physics control* (using the same text as in previous experiments), *narrative* (using the same four texts as in Experiment 3), *weak argument* (using the same four arguments as in Experiment 3), and two new conditions. The first new condition was a 404 word *charity control* condition, describing founding dates and other dry facts about five charities (none included among the donation options later in the experiment). The aim was to make charitable donation salient without offering either an engaging narrative or an argument for charitable giving. The second new condition was a *strong argument* condition, using four arguments (length 176 to 491 words) found to be effective in increasing donations to charity in unpublished earlier research employing a design similar to Experiments 2 and 3 (Schwitzgebel & Cushman 2020). These four arguments had little in common apart from being the winning arguments in a contest that

tested twenty arguments, submitted by blog readers, for their effectiveness in increasing rates of donation relative to other arguments and a control condition. Our expectation was that the arguments that performed well in that contest might be more effective than the arguments used in Experiments 1-3a. The aim was to see if these arguments could perform as well as the narratives.

After reading their assigned text, participants saw the same attitude questions and donation opportunity as in Experiments 2 and 3. Half of the participants viewed the attitude questions before the donation opportunity, and half viewed the donation opportunity first. After the same narrative transport questionnaire as in Experiments 1-3, participants were asked whether the text was "more like a narrative story" or "more like an argument with a conclusion". The questionnaire concluded with questions about gender, age, and political attitude. In this study, participants weren't asked if they would like to share their email address with a charitable organization, given the low rates of answering that optional question in Experiments 2-3.

5.2 Results

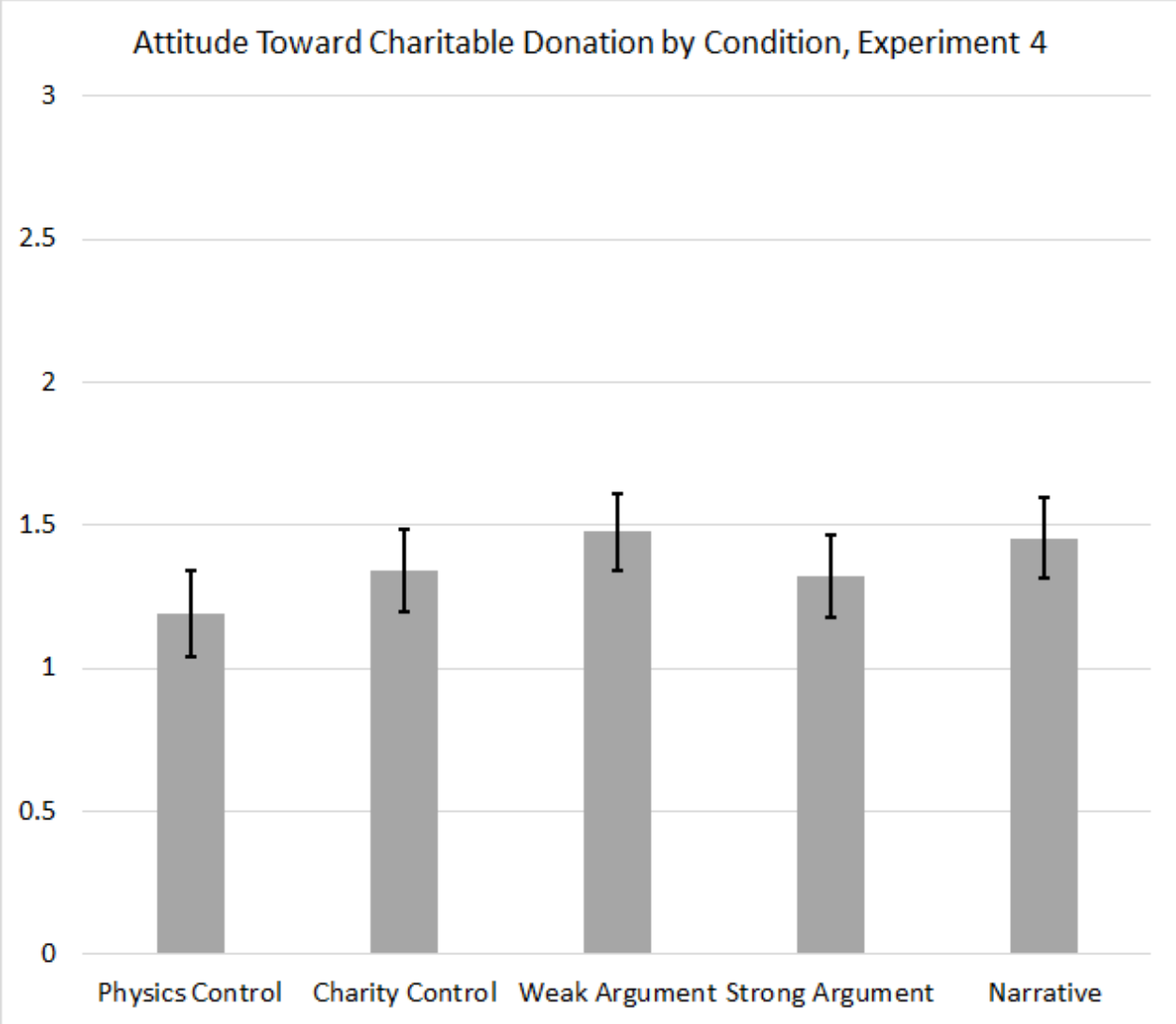
Attitude scale. The five questions constituting the attitude scale again showed good intercorrelation, with a Cronbach's alpha of .87, a single factor explaining 68% of the variance in a Principal Component Analysis, and a minimum factor loading of .74. Replicating Experiments 1-3, attitude correlated well with donation, $r = .47$ ($p < .001$).

Attitude by condition. Figure 8 shows attitude by condition. As in previous experiments, expressed attitude varied significantly by condition, with means of 1.19 in the physics control, 1.34 in the charity control, 1.48 in the "weak argument", 1.42 in the "strong argument", and 1.45 in the narrative condition (pooled $SD = 1.23$, $F[4,1427] = 2.47$, $p = .043$). However, contrary to

our expectations, the only statistically detectable difference between conditions in a Tukey post-hoc test was weak argument versus physics control. Pre-registered one-tailed t-tests comparing the narrative conditions and each argument condition against the combined control conditions showed small but statistically significant differences in the predicted direction for both the weak arguments and the narratives (M_{WA} = 1.48, M_C = 1.27, pooled SD = 1.22, $t(850) = 2.38$, $p = .009$; M_N = 1.45, M_C = 1.27, pooled SD = 1.24, $t(844) = 2.08$, $p = .019$; M_{SA} = 1.32, M_C = 1.27, pooled SD = 1.26, $t(856) = 0.61$, $p = .27$).

It is unlikely that the difference between these results and those of previous experiments was due to the fact that some participants received the donation question before the attitude questions. Mean attitude and mean donation did not detectably vary by order of presentation (attitude: M_{1st} = 1.35, M_{2nd} = 1.37, SD = 1.23, $t(1430) = 0.38$, $p = .70$; donation M_{1st} = \$4.23, M_{2nd} = \$4.44, SD = \$3.34, $t(1430) = 1.19$, $p = .23$), and in both orders of presentation the weak argument and the narrative conditions performed similarly. In Sections 5.3 and 5.4, we will address possible interpretations of the difference between these results and our expected results.

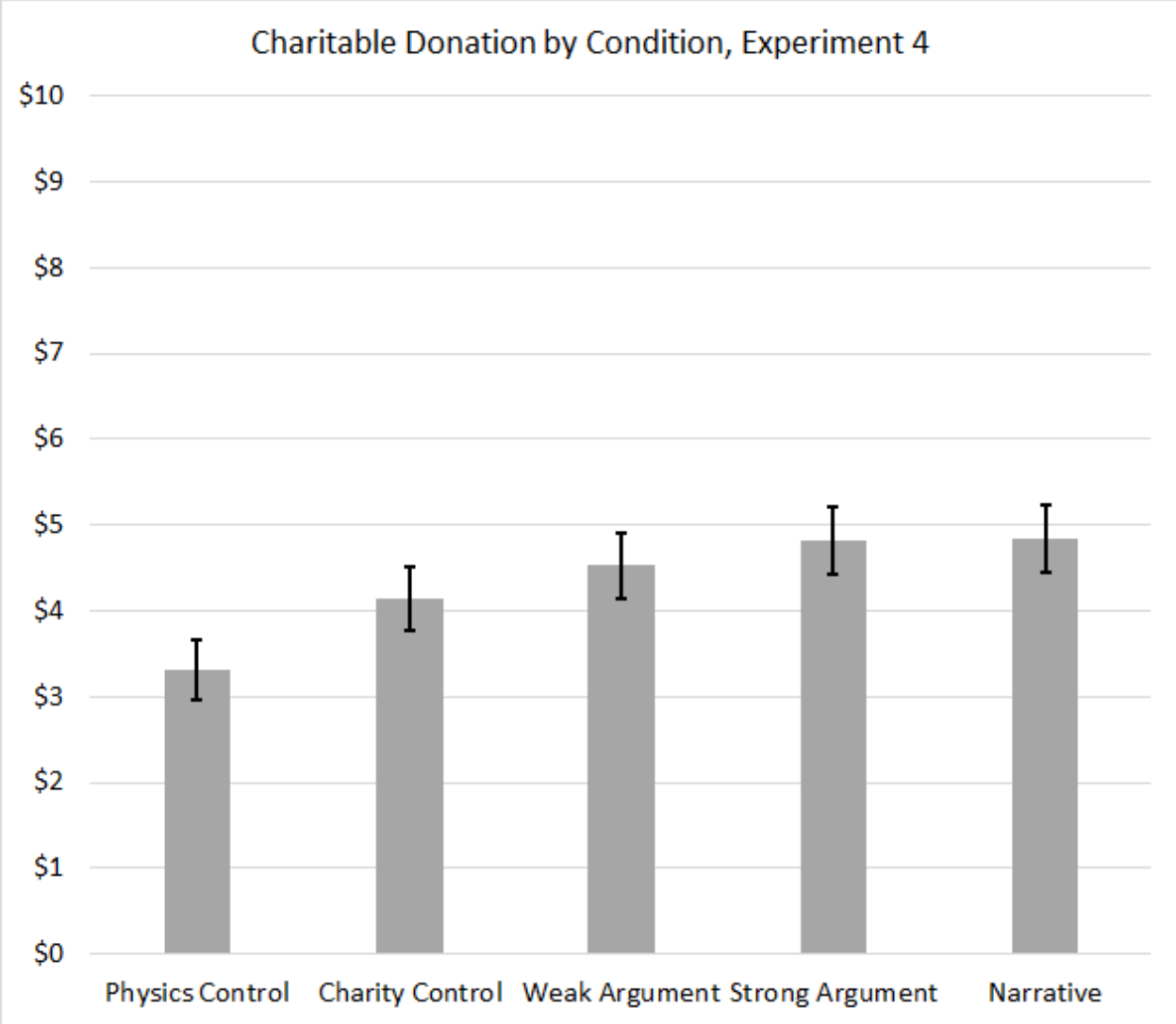
Figure 8: Mean attitude toward charitable giving by condition, on a scale from -3 to +3, Experiment 4. Error bars represent 95% confidence intervals.



Donation by condition. Figure 9 shows donation amount by condition. Again, the results were somewhat different than expected. Donations were lower in the physics control condition than in any other condition, with a mean of $3.31 in the physics control condition, $4.15 in the charity control condition, $4.53 in the weak argument condition, $4.81 in the strong argument condition, and $4.84 in the narrative condition (pooled *SD* = 3.29, $F[4,1427] = 10.51$, $p < .001$; Tukey post-hoc: all conditions > PC). Preregistered one-tailed *t* tests found donations higher in the narrative and both argument conditions than in the combined control conditions (*p*'s < .001), but given the different performance of the two control conditions a more conservative

comparison focuses on narrative and each argument condition compared to the charity control. We found donation to be statistically higher in the narrative and strong argument conditions than in the charity control, while the weak arguments showed a statistically marginal trend (M_N = \$4.84, M_{CC} = \$4.15, pooled SD = 3.31, $t(564) = 2.49$, $p = .007$; M_{SA} = \$4.81, M_{CC} = \$4.15, pooled SD = 3.32, $t(576) = 2.40$, $p = .008$; M_{WA} = \$4.53, M_{CC} = \$4.15, pooled SD = 3.27, $t(570) = 1.38$, $p = .083$). Patterns were similar whether the donation question was asked before or after the attitude questions.

Figure 9: Mean donation of a 10% chance of \$10 by condition, Experiment 4. Error bars represent 95% confidence intervals.



Differences among narratives and among arguments. As in Experiment 3, the four narratives did not perform detectably differently from each other, on either attitude or donation. Neither did the arguments perform detectably differently, despite our expectation that some would be "weak" and others "strong". However, power was somewhat limited. For donation by argument, average donation rates ranged from $4.30 (for one of the "weak" arguments) to $5.23 (for one of the "strong" arguments) (pooled SD = 3.42, F[7,578] = 0.53, p = .81).

Manipulation check. Fully 99% of participants described the narrative condition as "more like a narrative story" than "more like an argument with a conclusion". Conversely, the arguments were described as more like arguments (83% for the weak arguments and 75% for the strong arguments). Given the forced choice, participants in the charity control condition tended to describe it as more like a narrative (92%) while they were split in the physics control condition (54% narrative; overall $\chi^2[4] = 658.7$, $p < .001$). Among the eight arguments, two appeared to be experienced as having more narrative elements than the others (arguments 5 and 9, see Supplementary Online Materials for the text of these arguments) with 34% and 31% of participants rating them as more like a narrative, while the remaining six arguments were rated as more like a narrative by only 16-18% of participants ($\chi^2[7] = 16.3$, $p = .022$).

Transport. The narrative transport results were similar to Experiments 1-3, except perhaps with slightly lower means in all conditions (Cronbach's alpha = .79; $M_{PC} = 0.18$, $M_{CC} = 0.14$, $M_{WA} = 0.51$, $M_{SA} = 0.55$, $M_N = 0.75$ (pooled $SD = 0.85$, $F[4, 1427] = 26.1$, $p < .001$; Tukey post-hoc all three interventions higher than both controls and the narrative condition higher than both of the argument conditions). Correlations between transport and donation appeared to be lower in the physics control ($r = .16$, $p = .07$) than in the four other conditions (r's .28 to .40, p's < .001)). The relatively high correlation between transport and donation for the charity control condition ($r = .31$) might reflect that some participants were more transported and influenced by the text than we had expected (which might in turn explain the higher donations in that condition than in the physics control condition); or alternatively, participants who already have positive inclinations toward charitable giving might be disposed to express more positive attitudes toward texts describing dry facts about charities, creating the misleading appearance of transportation.

Regression and mediation analyses. We created regression analyses similar to those in Experiments 1-3, described in more detail in Appendix A. As in Experiments 1-3, transport mediated both attitude and donation. We detected no direct effect of assignment to the narrative condition on donation and, unexpectedly, a small *negative* direct effect on attitude.

5.3 Discussion

Experiment 4, with an improved design, partly confirmed the results of Experiments 1-3a. As in previous experiments, participants exposed to the narratives donated more than participants exposed to the physics text control condition.

However, the results differed from and expanded upon Experiments 1-3a in a few important respects. First, in contrast with Experiments 1-3a, participants exposed to the "weak" arguments, which we expected to be ineffective, expressed more positive attitudes toward charitable giving and donated more to charity than those exposed to the physics control condition. Participants who viewed new arguments drawn from Schwitzgebel & Cushman (2020), which we expected to be effective, also gave more to charity than those in the physics control condition. But contrary to our expectations we did not detect a difference between the "strong" arguments and the "weak" arguments. Participants exposed to the new charity control condition gave more than those in the physics control condition, despite low narrative transportation.

Evidence was mixed regarding effectiveness of narrative or argument relative to the charity control. The differences were not always statistically significant in Tukey tests, but participants in the narrative and strong argument conditions did donate significantly more than those in the charity control condition as measured by pairwise t-tests. We speculate that

compared to the physics control, the charity control might have made charity more salient, increased experimenter demand, reminded people of the importance and international reach of the major charities described, or implicitly communicated that many people donate to charity. Any of these factors might explain the higher rates of charitable giving in the charity control condition than in the physics control condition and the smaller difference between the charity control condition and the argument and narrative conditions.

Why was this replication only partial? We note two possibly important differences between Experiments 1-3a and Experiment 4 apart from intentional features of the design. First, we had moved from MTurk to Prolific. Perhaps the participant pool made a difference in this case. Second, the news cycle was dominated by the Russian invasion of Ukraine two weeks prior to the collection dates of March 8-9, 2022. Normally, we would not expect news events to have a big impact on attitudes toward donation. However, it's possible that world events influence participants' charitable attitudes and donation. For example, Schwitzgebel and Cushman (2020) found a marked decline in donation rates in a similar study of charitable giving in April, 2020, with many participants reporting economic hardship due to the onset of the COVID-19 pandemic. One post-hoc conjecture consistent with our results would be that with Ukraine in the news, participants might be slightly desensitized to narratives of suffering in foreign lands but also slightly more receptive to arguments for charitable giving.

However, another possible explanation of the difference is chance variation, which we will now explore.

5.4 Power Analysis and Meta-Analysis

The most comparable conditions between experiments are the physics control, the narrative condition, and the weak argument condition, which are similar across Experiments 2, 3, 3a, and 4 for donation and across all experiments for attitude (apart from the addition of alternative narratives and alternative weak arguments starting in Experiment 3). Table 2 shows mean donations and mean attitude scores for all five experiments, for these three conditions.

TABLE 2: Mean donations and mean attitude scores for all five experiments

Condition	Experiment	Mean donation	Mean attitude score
Physics control	1	–	1.38
	2	$3.97	1.20
	3	$3.39	1.04
	3a	$3.21	1.08
	4	$3.31	1.19
Weak argument	1	–	1.35
	2	$4.17	1.16
	3	$3.88	1.23
	3a	$3.84	1.24
	4	$4.53	1.48
Narrative	1	–	1.79
	2	$4.70	1.62
	3	$4.55	1.44
	3a	$4.29	1.50
	4	$4.84	1.45

As is evident from the table, the overall results are broadly consistent with donation rates across conditions of about $3.50 for the physics control, $4.00 for the weak argument, and $4.50 for the narrative, and attitude scores of about 1.10 for the physics control, 1.25 for the weak argument, and 1.50 for the narratives.

Given standard deviations of about $3.20 for donation and about 1.25 for attitude score, the relevant effect sizes to differentiate the weak argument from the control condition and the narrative condition are in the ballpark of d = .15 for both attitude and donation. We would thus need sample sizes of about 550 per condition for an 80% chance of detecting an effect in a one-tailed *t*-test (and larger samples for ANOVAs with Tukeys). With the sample sizes actually employed, ranging from about 200-300, post-hoc analysis suggests only about a 50% chance of detecting an effect, suggesting insufficient power. In contrast, for an effect size of d = .30, the approximate effect size when comparing narrative and control, samples of 200-300 are over 90% likely to detect an effect. We thus conjecture that our sample sizes were sufficient to *usually* detect the difference between narrative condition and physics control but only *sometimes* detect the difference between the weak argument condition and the other two conditions.

Given the high similarity of design and dependent measures across experiments, we see two ways to test the hypothesis that narrative is more effective than weak argument which is more effective than physics control, with small effect sizes. One approach is simply to pool the attitude data across all experiments and the donation data across Experiments 2, 3, 3a, and 4. Using this approach, the mean donation in the control condition across experiments was $3.46 [$N$ = 977], statistically less than the mean donation of $4.11 [$N$ = 1052] in the weak argument conditions (pooled SD = 3.11, $t(2027) = 4.70$, $p < .001$), which was in turn statistically less than the mean donation of $4.62 [$N$ = 913] in the narrative conditions (pooled SD = 3.27, t(1963) =

3.42, $p = .001$). Similarly, the mean attitude score in the control condition across experiments was 1.18 [$N = 1231$], statistically less than the mean attitude score of 1.28 [$N = 1347$] in the weak argument conditions (pooled $SD = 1.21$, $t(2576) = 2.19$, $p = .029$), which was in turn statistically less than the mean attitude score of 1.56 [$N = 1163$] in the narrative conditions (pooled $SD = 1.19$, $t(2508) = 5.80$, $p < .001$). Nonetheless, stepwise regressions find some main effects for experiment number as well as some condition-by-experiment-number interactions, so it remains likely that the differences between the experiments were not wholly due to chance.

Another approach to exploring the possibility of small N < A < C effects, which does not require pooling data across experiments, employs a meta-analytic procedure described in Rosenthal & Rosnow (1991), combining *df*-weighted *z*-scores across the studies (with no file-drawer adjustment, since there are no unreported versions of this experiment). Using this method, overall, participants exposed to the physics control text donated significantly less than those exposed to the weak arguments ($z = 4.64$, $p < .001$), who in turn donated significantly less than those exposed to the narratives ($z = 3.34$, $p < .001$). Similarly, participants exposed to the physics control text expressed significantly less positive attitudes toward charitable giving than those exposed to the weak arguments ($z = 2.21$, $p = .013$), who in turn expressed significantly less positive attitudes than those exposed to the narratives ($z = 5.69$, $p < .001$).

6. General Discussion

Our four experiments aimed to investigate the impact of ethical arguments versus narratives on charitable giving. We wanted to assess whether arguments of the sort that philosophers tend to give could be as effective as narratives at boosting charitable donations, and if so why.

Effects of narrative. We exposed participants to one of four emotionally engaging narratives about a child or teen in a developing country whose life was dramatically improved by a charitable organization. Participants in this condition consistently expressed slightly more positive attitudes toward charitable giving than did participants in a control condition, and they also gave approximately $1 more out of $10, in either a hypothetical donation (Experiment 1) or with a 10% chance of receiving the $10 (Experiments 2-4). These results extend research which has shown that descriptions of particular identifiable individuals can motivate charitable giving (e.g., Lee & Feeley 2016) and that narratives can influence attitudes and behavior on other contemporary moral issues (e.g., Paluck 2009).

Effects of argument. In the same four experiments, participants read one of several philosophical arguments for charitable giving. In some cases participants exposed to the philosophical argument did give slightly more than participants in a control condition. Positive attitudes toward charitable giving sometimes differed between the argument and control conditions, but with less consistency and apparently less force than the narratives.

In other words, while exposure to narrative had a substantial and consistent effect on attitude and donation, exposure to philosophical argument had a smaller and less consistent effect. These results comport well with prior research that has found that philosophical arguments can sometimes increase charitable donations relative to control conditions (Lindauer et al. 2020; Schwitzgebel & Cushman 2020; Buckland et al. 2021).

Photos. Adding a photo of the actual beneficiaries of charitable giving – the real-life protagonists of the narratives – had no detectable effect on attitude or donation. This null result was surprising given prior research reporting that photos of beneficiaries can increase charitable donations (e.g., Bergh & Reinstein 2021). However, rich narratives or arguments might be

sufficient for readers to imaginatively identify particular victims in need, making images superfluous.

Narrative transport. Why do narratives and arguments increase charitable donations when they do? We hypothesized that narratives move people to donate more in part by making readers engage more with the message of the text. In all four experiments, aggregated narrative transport predicted both expressed attitude and donation (hypothetical or lottery). Transportation consistently mediated both attitude and donation, with little or no statistically detectable direct effect from exposure to narrative once transport was taken into account. These results suggest that appeals to donate can be effective whether or not they are "emotional vs. rational" or "positive vs. negative" (compare Erlandsson et al. 2016; Erlandsson et al. 2018; Lindauer et al. 2021). Across all of our studies, what appeared to matter most was the extent to which participants found the text *engaging* – at least as measured by questions like "I wanted to learn how the text ended", "While reading the text, I could easily picture the events in it taking place", and "The text affected me emotionally." Importantly, this is but one operationalization of "narrative" and does not map perfectly onto all conceptions of narrative understanding (Currie 2010).

Limitations and future directions. The behavior of online research participants might differ from the behavior of others not in an online research setting, and behavior with small amounts of money might not reflect what people would do with larger amounts. Our results might underestimate willingness to donate, since some participants might have been motivated to donate but to a charity of their choice, outside of the experiment, rather than to one of the charities included in our study. Effects might have been short term and subject to social desirability bias and experimenter demand. The range of narratives and arguments was limited

(primarily to positive rather than negative appeals). In future work, by systematically varying features of the narratives and features of the arguments, we hope to further explore what aspects of narrative and argument are and are not effective in changing people's attitudes and behavior. It might also be possible to explore whether people with different reasoning styles are differently influenced by narratives versus arguments (compare with individual differences studies on the identifiable victim effect, e.g., Friedrich & McGuire 2010).

Philosophical implications. Is philosophical reasoning valuable, and why? We cannot fully address this question here. We would need to know much more about the psychological effects of philosophical reasoning, including under what conditions reading or thinking about philosophical issues influences one's attitudes and behavior. In practice, almost everyone in our society, even outside the classroom, has some exposure to philosophical arguments for charitable giving and some exposure to emotionally moving narratives about people rescued by charitable donations. For this reason, we expect that any effect on attitudes toward charitable giving from exposure to a brief argument or narrative would be temporary for most participants and probably not large in size. In this respect, brief online studies more closely resemble encountering a plea from a charitable organization or an argument advanced in a short conversation with a stranger than they will resemble extended exposure to detailed philosophical arguments (typically mixed with thought experiments containing narrative elements) that one would encounter in extended classroom treatment or in reading a monograph on the topic.

We did find, nevertheless, that some philosophical arguments can have an effect on ordinary people's attitudes and behavior, perhaps especially when the arguments create more narrative transport. If part of the aim of philosophical argumentation is to sway hearers' attitudes and behavior, philosophers should be aware that arguments with narrative elements might be

more effective than dry, didactic presentations. From one perspective, this might seem obvious. However, in light of research suggesting that studying ethical arguments is motivationally impotent (Schwitzgebel & Rust 2014), it might be reassuring that arguments about charitable giving can influence behavior, if they are appropriately engaging.

References

Abend, G. (2014). *The moral background.* Princeton: Princeton.

Aristotle. 4th c. BCE/1962. *Nicomachean ethics*, trans. M. Ostwald. New York: Macmillan.

Bekkers, R. & Wiepking, P. (2011). A literature review of empirical studies of philanthropy: Eight mechanisms that drive charitable giving. *Nonprofit and Voluntary Sector Quarterly, 40*, 924-973.

Bergh, R., & Reinstein, D. (2021). Empathic and numerate giving: The joint effects of victim images and charity evaluations. *Social Psychological and Personality Science* 12(3): 407-416.

Bhati, A., & Hansen, R. (2020). A literature review of experimental studies in fundraising. *Journal of Behavioral Public Administration, 3*, 1-19.

Buckland, L., Lindauer, M., Rodríguez-Arias, D. & Véliz, C. (2021). Testing the Motivational Strength of Positive and Negative Duty Arguments Regarding Global Poverty. *Review of Philosophy & Psychology:* 1–19. Doi:10.1007/s13164-021-00555-4.

Chang, C. (2009). "Being Hooked" by editorial content: The implications for processing narrative advertising. *Journal of Advertising, 38*, 21-33.

Currie, G. (2010). *Narratives and narrators: A philosophy of stories*. Oxford University Press.

Erlandsson, A., Nilsson, A., & Västfjäll, D. (2018). Attitudes and donation behavior when reading positive and negative charity appeals. *Journal of Nonprofit & Public Sector Marketing*, *30*, 444-474.

Erlandsson, A., Västfjäll, D., Sundfelt, O., & Slovic, P. (2016). Argument-inconsistency in charity appeals: Statistical information about the scope of the problem decrease helping

toward a single identified victim but not helping toward many non-identified victims in a refugee crisis context. *Journal of Economic Psychology* 56: 126-140.

Escalas, J. E., Moore, M. C., & and Britton, J. E. (2004). Fishing for Feelings? Hooking Viewers Helps! *Journal of Consumer Psychology, 14*, 105-114.

Friedrich, J., & McGuire, A. (2010). Individual differences in reasoning style as a moderator of the identifiable victim effect. *Social Influence, 5* (3), 182-201

Genevsky, A., Västfjäll, D., Slovic, P., & Knutson, B. (2013). Neural underpinnings of the identifiable victim effect: Affect shifts preferences for giving. *Journal of Neuroscience, 23*, 17188-17196.

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology, 79*, 701-721.

Green, M. C., & Sestir, M. (2017). Transportation theory. In P. Rössler, ed., *International Encyclopedia of Media Effects.* New York: Wiley Blackwell.

Han, H., Workman, C. I., May, J., Scholtens, P. Dawson, K. J., Glenn, A. L., Meindl, P. (2022). Which Moral Exemplars Inspire Prosociality? *Philosophical Psychology*. DOI: 10.1080/09515089.2022.2035343

Hansson, S. O. (2020). Do moral philosophers have to be moral? *Theoria, 86,* 433-438.

Hart, P. S., Lane, D. and Chinn, S. (2018). The elusive power of the individual victim: Failure to find a difference in the effectiveness of charitable appeals focused on one compared to many victims. *PLoS ONE 13*(7): e0199535.

Hsu, T. C. (2007). *CPO physical science, student ed.* Nashua, NH: CPO Science.

Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind. Science, 342 (6156), 377-380.

Kogut, T., & Ritov, I. (2011). The identifiable victim effect: Causes and boundary conditions. In D. M. Oppenheimer and C. Y. Olivola (eds.), *The science of giving.* New York: Psychology Press (pp. 133-145).

Koopmann, E. M., & Hakemulder, F. (2015). Effects of literature on empathy and self-reflection: A theoretical-empirical framework. *Journal of Literary Theory, 9,* 79-111.

Lamarque, P. (2004). On not expecting too much from narrative. *Mind & Language* 19(4): 393-408.

Lee, S., & Feeley, T.H. (2016). The identifiable victim effect: A meta-analytic review. *Social Influence, 11,* 199-215.

Lindauer, M., Mayorga, M., Greene, J., Slovic, P., Västfjäll, D., & Singer, P. (2020). Comparing the effect of rational and emotional appeals on donation behavior. *Judgment and Decision Making, 15,* 413–420.

MacAskill, W. (2015). *Doing good better: How effective altruism can help you make a difference*. Avery.

Nussbaum, M. C. (1995). *Poetic justice.* Boston: Beacon.

Paluck, E. L. (2009). Reducing Intergroup Prejudice and Conflict Using the Media: A Field Experiment in Rwanda. *Journal of Personality and Social Psychology* 96: 574–587.

Rust, J., & Schwitzgebel, E. (2014). The moral behavior of ethicists and the role of the philosopher. In H. Rusch, M. Uhl, and C. Luetge, eds., *Experimental Ethics.* New York: Palgrave.

Schönegger, P., & Wagner, J. (2019). The moral behavior of ethics professors: A replication-extension in German-speaking countries. *Philosophical Psychology, 32,* 532-559.

Schwitzgebel, E., Cokelet, B., & Singer, P. (2020). Do ethics classes influence student behavior? Case study: Teaching the ethics of eating meat. *Cognition*, 203: 104397.

Schwitzgebel, E., & Cushman, F. (2020). Contest winner! A philosophical argument that effectively convinces research participants to donate to charity. Blog post at The Splintered Mind (May 23), URL: http://schwitzsplinters.blogspot.com/2020/06/contest-winner-philosophical-argument.html

Schwitzgebel, E., & McVey, C. (2018). Narrative but not philosophical argument motivates giving to charity. Blog post at The Splintered Mind (Nov. 21), URL: http://schwitzsplinters.blogspot.com/2018/11/narrative-but-not-philosophical.html

Schwitzgebel, E., & Rust, J. (2014). The moral behavior of ethics professors: Relationships among self-reported behavior, expressed normative attitude, and directly observed behavior. *Philosophical Psychology, 27,* 293-327.

Schwitzgebel, E., & Rust, J. (2016). The moral behavior of ethicists. In J. Sytsma and W. Buckwalter, eds., *Companion to experimental philosophy.* New York: Wiley Blackwell.

Singhal, A., & Rogers, E. M. (1999). *Entertainment-education: A communication strategy for social change.* Yahweh, NJ: Lawrence Erlbaum.

Small, D. A., Loewenstein, G., & Slovic, P. (2007). Sympathy and callousness: The impact of deliberative thought on donations to identifiable and statistical victims. *Organizational Behavior and Human Decision Processes, 102,* 143-153.

Small, D. A., & Verrochi, N. M. (2009). The face of need: Facial emotion expression on charity advertisements. *Journal of Marketing Research, 46,* 777-787.

Velleman, J. D. (2003). Narrative explanation. *Philosophical Review 112,* 1-25.