Philosophical Arguments Can Boost Charitable Giving

Kirstan Brodie[a], Eric Schwitzgebel[b], Jason Nemirow[c], and Fiery Cushman[c]

[a] Department of Psychology, Cornell University, 211 Uris Hall, Ithaca, NY 14850, U.S.A.

Email: klb322@cornell.edu

[b] Department of Philosophy, University of California at Riverside, 900 University Ave, Riverside, CA 92521, U.S.A. Email: eschwitz@ucr.edu

[c] Department of Psychology, Harvard University, 33 Kirkland St, Cambridge, MA 02138, U.S.A.

Emails: nemirow@fas.harvard.edu, cushman@fas.harvard.edu

Author Note

Correspondence concerning this article should be addressed to Kirstan Brodie, Department of Psychology, Cornell University, 211 Uris Hall, Ithaca, NY 14850, U.S.A. Email: klb322@cornell.edu.
Declaration of Interest: The authors declare no conflicts of interest. Funding: This research was supported by the John Templeton Foundation (Grant No. 61061) and the National Science Foundation Graduate Research Fellowship (Grant No. DGE–2139899). Open Science: All study materials, data, and analysis scripts are publicly available (https://osf.io/ujzfc/?view_only=aa1747631e7b4350b8c5203ad6655642). Acknowledgements: The authors thank Christopher McVey for valuable feedback, and Isobel Munday for assistance with data collection.

Abstract

Do reasoned arguments increase charitable giving? Evidence is mixed. One possible explanation is that arguments vary in effectiveness. We crowdsourced 90 philosophical arguments in favor of charitable giving, coding them for a variety of features. In Study 1 (N=2161, MTurk), participants read either a control text or one of five arguments preselected as especially promising. Participants who read the philosophical arguments expressed more positive attitudes towards donating and donated about $1 more on average of a potential $10 bonus than participants who read the control text. Study 2 (N=8982, MTurk) employed all 90 arguments, with each participant reading one argument. We analyzed the relationship between the coded argument features and donation behavior, and found arguments to be more effective when they mentioned children, mentioned the specific impacts of donating, and/or emphasized that large benefits can be achieved at low cost. Among the features less predictive of donation: addressing counterarguments, mentioning the equivalence of persons or the participant's own assumed good fortune, appealing to religion or expert authorities, or appealing to the participant's self-interest. Together, these studies provide evidence that reasoned arguments can indeed shift moral behavior such as charitable giving, but that some arguments are simply more effective than others. Further, we identify specific argument features that are most and least effective in eliciting donations.

Keywords: charitable giving, reasoning, experimental philosophy, morality, persuasion.

Philosophical Arguments Can Boost Charitable Giving

People often use reasoned arguments to try to change each other's moral attitudes and behavior. It is unclear, however, how well this works (Haidt, 2001; May, 2018; Mercier, 2011). Philosophical arguments present an especially powerful test of this question. Philosophers are trained experts in reasoned analysis and therefore, if reason is persuasive, then philosophical arguments might be especially potent. On the other hand, if reasoned argument is generally unpersuasive, then philosophical arguments should be impotent to change attitudes or behavior, unless they incorporate other effective elements.

Existing work provides some evidence in favor of each of these possibilities. On the one hand, philosophers who specialize in ethics are no more likely to perform commonplace praiseworthy actions—such as donate to charity, vote, clean up litter, donate blood, or avoid joining the Nazi party in 1930s Germany—than comparison groups of professors specializing in other areas of philosophy or in other departments at the same universities (Hou et al., 2024; Schönegger & Wagner, 2019; Schwitzgebel, 2019; Schwitzgebel & Rust, 2016). On the assumption that philosophers who specialize in ethics are especially exposed to reasoned arguments in favor of virtuous behaviors, these results suggest that such arguments are inert.

On the other hand, professional philosophers are an unusual group, and several recent studies have found that exposing college students to philosophical arguments in favor of vegetarianism led to a significant reduction in the purchase of meat products from campus vendors over subsequent weeks or months (Jalil et al., 2020, 2023; Schwitzgebel et al., 2020, 2023). In these studies, however, the reasoned arguments were paired with other features such as emotionally evocative descriptions of factory farming, social support from authority figures in the classroom, health information, pledge opportunities, etc. Thus, the distinctive role of reasoned argument is unclear. Other studies show that exposure to reasoned argument can change participants' moral judgments, but leave open the question of whether these arguments change their behavior (Nichols & Knobe, 2007; Paxton & Greene, 2010).

A number of studies specifically explore whether philosophical arguments in favor of charitable giving can persuade people to donate some of a surprise bonus award to charity, above the amount

donated by those in a control condition (Buckland et al., 2022; Grodeck & Schoenegger, 2023; Lindauer et al., 2020; Schwitzgebel et al., 2022). Evidence for this effect is mixed and, in some conditions, a philosophical argument was no more effective than a control text in eliciting donations. Given this heterogeneity, a key question is whether certain features of philosophical arguments work, and others do not.

In the present study we therefore crowdsourced the generation of nearly 100 reasoned, philosophical arguments in favor of charitable giving and then tested their effects on behavior. This allows us to explore not only whether reasoned arguments can change behavior, but also what features of those arguments are effective or ineffective. To this end, we held a contest: entrants were asked to write a philosophical argument in favor of donating money to charity, and the winning entry would receive $500 in prize money as well as a $500 donation on their behalf to a charity of their choice.

In order to explore the distinctive role of rational persuasion, we asked entrants to restrict the contents of their argument to a rather narrow category. Previous research has demonstrated that narrative transport, emotionally arousing content, and mentions of "identifiable victims" are all particularly effective in eliciting donations (Bagozzi & Moore, 1994; Erlandsson et al., 2018; Jenni & Loewenstein, 1997; Kogut & Ritov, 2011; Schwitzgebel et al., 2022). We sought to test whether philosophical arguments can effectively persuade people to donate without relying on such features. Accordingly, contest entrants were explicitly instructed to avoid narrative elements (e.g., descriptions of specific events in history or in individuals' lives); emotional rhetoric (e.g., vivid descriptions of suffering), and mentions of specific individuals (i.e. identifiable victims), aside from the reader themselves.

Our investigation was divided into two parts. First, we asked whether any philosophical arguments could reliably increase charitable giving above a neutral baseline. To do this, from among the submitted arguments we selected 20 that seemed especially promising, adhered closely to our criteria, and represented a range of distinct argument types. We tested their effectiveness in eliciting donations from participants on Amazon Mechanical Turk, identifying the most effective five among this set. We then tested a new set of participants either with one of these especially strong arguments, or with a control text

excerpted from a middle school physics textbook, to assess whether these arguments elicited higher-than-baseline levels of giving.

In the second part of our investigation, we tested the effectiveness of all 90 submitted arguments in increasing charitable giving, with the aim of discovering what types of arguments are relatively effective and ineffective. By coding the arguments for various features – such as whether they mention specific hardships that can be addressed through charitable giving, explicitly rebut counterarguments against charitable giving, or appeal to the participant's self-interest – we could explore the extent to which the presence of those features does or does not reliably encourage charitable giving.

1. Study 1: Can reasoned argument increase charitable giving?

1.1 Method

We received a total of 90 arguments following a solicitation on social media. For our first study, we hand-selected 20 that were heterogeneous in content, adhered closely to our guidelines, and seemed promising to us. We recruited 2532 participants from Amazon Mechanical Turk, excluding 89 who failed a comprehension check, leaving 2443 participants for analysis. Each participant was presented with one of these 20 arguments; thus, roughly 125 participants viewed each argument. Participants were then informed that they had a 10% chance of receiving a $10 bonus, and were given the opportunity to donate a portion of that possible bonus to one of six well-known, effective charities. From this dataset we then selected the five arguments that elicited the highest mean donations and submitted them to a second phase of testing. In this second round of testing, after 110 exclusions, approximately 335 participants viewed each argument, while a sixth group of 472 participants viewed a control text drawn from a middle school physics textbook (2161 participants total across all arguments and control). This group was used to establish a baseline measure of charitable giving in the absence of an argument. The sample sizes for this second phase of testing were determined via an *a priori* power analysis, and our final sample sizes provided 95% power to detect an effect size of $d = 0.3$ or greater in an independent-samples t-test with a 1% false-positive rate.

In both phases, following the donation question, participants were asked the extent to which they agreed with two statements: i) "It is morally good to give money to charities that help those in extreme poverty", and ii) "People like me should give money to charities that help people in extreme poverty" (-3, "Strongly disagree" to +3, "Strongly agree"). Finally, participants completed two additional exploratory measures: a 7-item Cognitive Reflection Test (adapted from Pennycook & Rand, 2019) and the 18-item Need for Cognition Scale (Cacioppo et al., 1984). Responses to both measures are included in our open dataset, but are not analyzed or discussed further.

All procedures reported in this manuscript were performed in compliance with relevant laws and was approved by the Committee on the Use of Human Subjects at Harvard University (Protocol No. IRB14-2016). All participants gave informed consent prior to participation, and their privacy rights have been observed. All studies, measures, manipulations, and participant exclusions are reported in the manuscript, and all study materials, data, and analysis scripts are publicly available (https://osf.io/ujzfc/?view_only=aa1747631e7b4350b8c5203ad6655642).

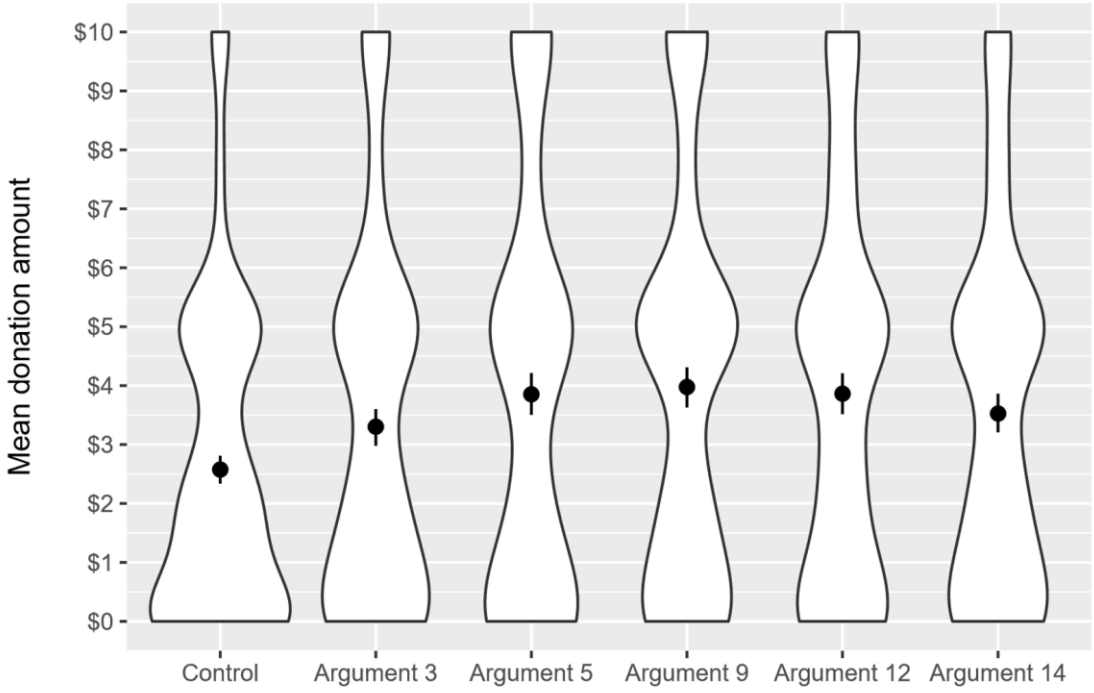
1.2 Results

We first sought to determine whether the 5 arguments identified in our two-phase procedure collectively outperformed the control text in eliciting donations, considering only the Phase 2 responses.

Each of the five arguments was compared pairwise with the control condition. Mean donation in the control condition was $2.58. All of the arguments performed about a dollar better and similarly, with means ranging from $3.30 for the lowest-performing argument to $3.98 for the best performing argument (SD_C = $2.70, SD_A = $3.05 to $3.43; ts $\geq$ 3.47 [equal variance not assumed], ps < .001; see Supplemental Tables 1 through 4 in Appendix F). In an analysis of variance (ANOVA) with just the five arguments, Argument 9 outperformed Argument 3 in a post-hoc Tukey pairwise comparison, but no other comparison was significant (see Supplemental Tables 5 and 6 in Appendix F).

Figure 1

Mean Donation by Type of Text Viewed



Note. Mean charitable donation by type of text viewed in Phase 2 of the contest. Error bars represent 95% confidence intervals.

The argument that elicited the highest mean donation, submitted by philosophers Matthew Lindauer and Peter Singer, won the contest. The text of Argument 9, the winning argument, follows:

> *Many people in poor countries suffer from a condition called trachoma. Trachoma is the major cause of preventable blindness in the world. Trachoma starts with bacteria that get in the eyes of children, especially children living in hot and dusty conditions where hygiene is poor. If not treated, a child with trachoma bacteria will begin to suffer from blurred vision and will gradually go blind, though this process may take many years. A very cheap treatment is available that cures the condition before blindness develops. As little as $25, donated to an effective agency, can prevent someone going blind later in life.*
>
> *How much would you pay to prevent your own child becoming blind? Most of us would pay $25,000, $250,000, or even more, if we could afford it. The suffering of children in poor countries*

must matter more than one-thousandth as much as the suffering of our own child. That's why it is good to support one of the effective agencies that are preventing blindness from trachoma, and need more donations to reach more people.

Participants' attitudes toward charitable giving — their agreement that donating is morally good and that people like themselves should donate — were also slightly more positive in the argument conditions than the control condition (morally good: M_C = 2.01, M_A = 2.14, pooled SD = 0.96, t[2159] = 2.49, p = .013; people like them: M_C = 1.33, M_A = 1.49, pooled SD = 1.35; t[2159] = 2.36, p = .019).

The goal of our contest was to generate philosophical arguments that might be effective in eliciting donations and to test their actual effectiveness compared to a control text. Indeed, the five arguments we tested all outperformed the control text, eliciting significantly higher mean donations. These preliminary results provide further evidence that philosophical arguments can sometimes persuade people to donate to charity. Given the mixed evidence for this effect in previous research, however, they also raise important questions. What makes some arguments more effective than others? Which argument features increase or decrease the amount people donate? The full set of entries to our contest — 90 philosophical arguments in total — provided us an opportunity to explore these questions.

2. Study 2: What features of philosophical arguments increase charitable giving?

To understand what makes philosophical arguments effective at persuading people to donate to charity, we employ our large, diverse set of arguments. After categorizing features of these arguments via experimenter coding and participant ratings, we then tested how these features predict charitable behavior (donation) and expressed charitable attitudes. To preview our results, the behavioral measure of charitable giving yielded several interesting and interpretable findings, while the attitudes measure showed smaller and less consistent effects. Therefore, in the following manuscript, we report the results of all preregistered results but place our primary focus on donation amounts. The research question, methods, and analysis plan for this study were preregistered (https://aspredicted.org/9mtf-tcmh.pdf).

2.1 Method

2.1.1 Participants

We recruited 9,041 participants from Amazon Mechanical Turk. (This number slightly exceeds the intended sample size of 9,000 listed in our preregistration; this over-recruitment was a result of a technical error on the MTurk platform. All data were collected prior to analysis.) This sample size was determined via an a priori power analysis: We used Monte Carlo simulations to assess parameter recoverability given this sample size and plausible assumptions about the structure of our data informed by prior research. These simulations indicated that if our data were generated from a set of feature weights that combine linearly to influence charitable behavior, the recovered feature weights (i.e., beta weights) from our statistical model would correlate at $0.5 < r < 0.6$ with the true feature weights. Of the 9,041 recruited participants, we excluded 59 whose open-ended comprehension check responses failed to meet our preregistered criteria, resulting in a final sample of 8,982 participants. We conducted a simulation-based sensitivity power analysis for each of the statistically significant predictors reported in our preregistered models (a total of 26 significant predictors across six linear mixed-effects models). We used the R package *simr* (Green & MacLeod, 2016) for these simulations, with our Study 2 data and each of our six models as inputs, 1000 simulations per analysis, and a 5% false-positive rate, to identify the power that our final sample had to detect the respective effects reported. The observed power for each of our reported effects ranged from 53% to 100%, with an average power across effects of 87.84% (*SD* = 12.90%). A table of the observed power for each individual effect is included in Supplemental Table 8 in Appendix F. All participants were United States residents and provided informed consent prior to participating.

2.1.2 Design & Materials

2.1.2.1 Philosophical Arguments. As described previously, we collected a total of 90 philosophical arguments via a contest whose entrants were asked to write an argument persuading the reader to donate to charity. All arguments were 500 words or less, and entrants were instructed to avoid using narrative elements, mentioning specific individuals, or including vivid emotional content. About half of the arguments were submitted by professional philosophers, psychologists, and experimental

economists, and about half were submitted by others who had heard about the contest through social media or other channels.

2.1.2.2 Questionnaire. Participants were each randomly assigned to view one of the 90 philosophical arguments. After reading the argument, they were then informed that they had a 10% chance of receiving an additional $10 bonus payment, and that they could elect to donate any whole-dollar amount (or none) of this potential bonus to a charity that they would select from a list of six "well-known, effective" charities. The amount of money they chose to donate, which we will refer to as "charitable behavior", was our primary dependent measure.

Following the donation question and charity selection, participants were asked to summarize the argument they previously read, which served as a comprehension check.

Next, participants were asked the extent to which they agreed with the following two statements, on 7-point scales ranging from "Strongly disagree" (-3) to "Strongly agree" (3):

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

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

Each participant's responses to these two questions were averaged to obtain a "charitable attitudes" score, which was a secondary dependent measure of interest.

2.1.2.3 Experimenter-rated Argument Features. In order to analyze which argument features would be most effective in soliciting donations, we coded the presence/absence of 20 features in each of the 90 arguments (e.g., "Does the argument mention religion?"; see Table 1 for the full list). We will refer to these as "experimenter-rated features". Two raters first independently coded the arguments, reaching 92% agreement, and then resolved coding discrepancies via discussion. Argument coding was completed prior to data collection. Among these 20 features, there were 8 which were present in more than 20% of the 90 arguments (and absent in the remainder), which we refer to as "high-frequency" features. The remaining 12 experimenter-rated features are referred to as "low-frequency" features. This categorization scheme, established in our preregistration plan, guides our interpretation of results. Because low-frequency features are present in only a small number of arguments, any empirical association with

charitable giving should be treated with some caution, given the possibility of confounding argument-specific effects (i.e., unmodeled variance at the level of particular arguments). We assign relatively greater weight to findings obtained with high-frequency features.

Table 1

Experimenter-rated Features

Feature	Description	Occurrences
Particular hardships	Mentions specific, physical needs or hardships, such as clean drinking water or blindness	49
Massive margins for poor	Appeals to the massive marginal utility of dollars transferred to poor recipients	37
Own good fortune	Refers to the reader's own assumed economic good fortune	35
Diminishing margins for rich	Appeals to the diminishing marginal utility of dollars kept by rich donors	28
Children	Mentions children	27
Experts	References experts (e.g., philosophers, scientists), either as a group or individually	27
Self-interest	Describes the possible intrinsic or extrinsic benefits of donating *for the donor*	26
Counterarguments	Mentions and addresses possible counterarguments	19
External sources	References external, material sources, such as empirical studies or news articles	18
Concrete benefit	Mentions a quantified, concrete benefit of donating (e.g., providing 3 mosquito nets)	17
Luck/randomness	Mentions luck or randomness as a determinant of life circumstances or events	14
Behavioral contagion	Mentions that donating might inspire other people to give or spark behavioral contagion	9
Religion	Appeals to religion, religious teachings/tenets, or religious authorities	9
Single hardship	Focuses on one, single need or hardship, such as trachoma	8
VOI, perspective taking	Appeals to veil-of-ignorance reasoning or other perspective-taking thought experiments	8
Common	Mentions that donation is a common behavior	7
Better than nothing	Mentions that giving something is better than giving nothing	6
Evolution	Mentions human evolutionary history	5
Equal value	Appeals to the notion of equality between individuals or groups	4
Particular near person	Mentions a particular person who is physically or emotionally near to the reader	3

Note. Experimenter-rated argument features. Gray and white color coding correspond to high- and low-frequency features, respectively.

2.1.2.4 Participant-rated Argument Features. Following the charitable attitudes questions, participants were asked to answer 16 additional questions about the philosophical argument they read. They provided ratings on 7-point scales, anchored on either end with opposing statements (e.g., "The reasoning in the text was flawed/ The reasoning in the text was sound"; "I did not find the text to be hopeful and encouraging/I found the text to be hopeful and encouraging"; see Table 2 for the full list). We used the answers to these questions, which we will refer to as "participant ratings", as additional "features" of the arguments to predict our dependent measures of interest: charitable behavior and charitable attitudes.

Table 2

Participant-rated Features

Feature	Description
Emotional	The text [left me cold/affected me emotionally].
Convincing	The text was [unconvincing/convincing].
Sound	The reasoning in the text was [flawed/sound].
Learned	[The text did not teach me anything new/I learned something new reading the text].
Will think	In the future, I [doubt I will remember much about this text/will probably think about this text again].
Agreed	I [disagreed/agreed] with what the text said.
Made a difference	The text [did not lead/led] me to offer more of my possible bonus money to charity than I would have if I hadn't read it.
Kept attention	[My mind wandered while reading the text/The text kept my attention well].
Easy	I found the text [difficult/easy] to understand.
Well-written	The text was [badly/well] written.
Vivid	I [did not/did] vividly imagine the events described in the text.
Hopeful	I [did not find/found] the text to be hopeful and encouraging.
Accusatory	I [did not find/found] the text to be accusatory or insulting.
Guilty	The text [did not make/made] me feel guilty.
Relevant	The text [is not/is] relevant to beliefs or goals that are important to me.
Shared	I [would not/would] like to see this text shared widely

Note. Participant-rated argument features.

2.2 Results

2.2.1 Charitable Behavior

2.2.1.1 The Range of Performance of Individual Arguments. See the online supplement for the full text and average performance of each of the 90 arguments. As noted there, the arguments performed significantly differently (ANOVA $F[89,8892] = 2.91$, pooled $SD = \$3.02$, $p < .001$), from a minimum average donation of $2.54 for Argument 19 to a maximum of $4.88 for Argument 31 (see Supplemental Table 7 in Appendix F).

2.2.1.2 Experimenter-rated Argument Features. As previously mentioned, we coded for the presence/absence of 20 different features in each of the 90 arguments. Features present in more than 20% of arguments were classified as high-frequency features. As outlined in our preregistered analysis plan, we ran a linear mixed-effects model to predict participants' donation amount (i.e. their charitable behavior) with the eight high-frequency features as fixed effects and argument index as a random effect. This model focused only on high-frequency features because, in an analysis of simulated datasets that guided our preregistered analysis plan, we found that the effects of low-frequency features were often poorly estimated and, furthermore, could contribute to the misestimation of the effects of high-frequency features. We used the package lmerTest (Kuznetsova et al., 2017) in R Statistical Software (v4.3.0; R Core Team, 2023) to estimate this model (and all regression models to follow). We found that mentioning children predicted a significant increase in donation amount ($b = 0.37$, $SE = 0.12$, $p = .001$) and addressing counterarguments predicted a significant decrease in donation amount ($b = -0.22$, $SE = 0.12$, $p = .044$).

Table 3

Model Summary of Linear Mixed-Effects Model Predicting Donation Amounts

Effect	Estimate	*SE*	95% CI		*p*
			LL	*UL*	
(Intercept)	3.19	0.09	3.02	3.36	<.001 ***
Counterarguments	-0.22	0.12	-0.45	0.01	.044 *
Self-interest	-0.14	0.10	-0.35	0.06	.147
Particular hardships	0.04	0.11	-0.18	0.26	.669
Experts	0.06	0.10	-0.15	0.26	.565
Own good fortune	0.06	0.11	-0.17	0.29	.583
Diminishing margins for rich	0.11	0.15	-0.18	0.40	.430
Massive margins for poor	0.23	0.13	-0.02	0.48	.054
Children	0.37	0.12	0.14	0.60	.001 **

Note. Results of linear model predicting donation amounts with high-frequency experimenter-rated argument features as fixed effects. P values are based on likelihood ratio tests as specified in the preregistration. *** $p < .001$, ** $p < .01$, * $p < .05$.

Next, as a test of the potential influence of low-frequency experimenter-rated features, we ran another linear mixed-effects model predicting participants' donation amount. We retained the two significant high-frequency features as fixed effects, and additionally included all 12 of the low-frequency features; the model was otherwise identical to the first. We found that only one of the original significant high-frequency predictors ("children", $b = 0.42$, $SE = 0.14$, $p = .002$) remained significant, while none of the remaining low-frequency predictors achieved significance.

Table 4

Model Summary of Linear Model Predicting Donation Amounts

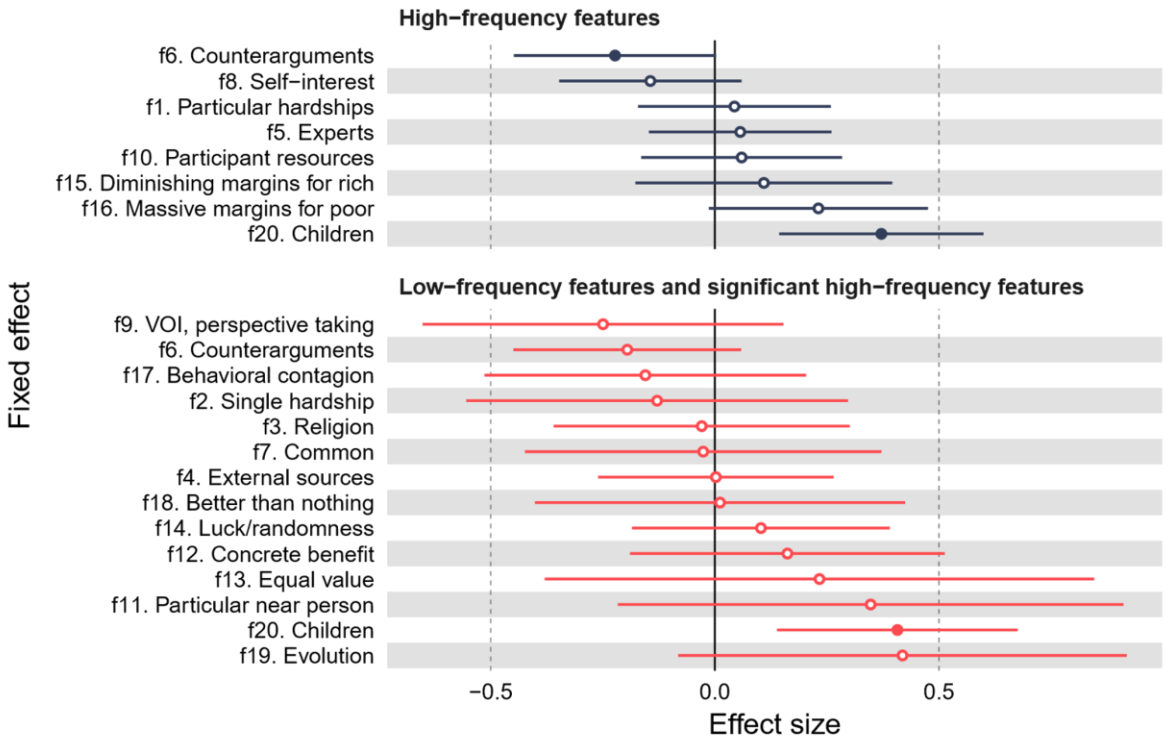
Effect	Estimate	*SE*	95% CI		*p*
			LL	*UL*	
(Intercept)	3.29	0.08	3.13	3.44	<.001***
VOI, perspective taking	-0.25	0.21	-0.66	0.16	.187
Counterarguments	-0.20	0.13	-0.45	0.06	.103
Behavioral contagion	-0.16	0.18	-0.52	0.21	.357
Single hardship	-0.13	0.22	-0.56	0.30	.514
Religion	-0.03	0.17	-0.37	0.31	.851
Common	-0.03	0.20	-0.43	0.38	.885
External sources	0.00	0.13	-0.27	0.27	.984
Better than nothing	0.01	0.21	-0.41	0.43	.957
Luck/randomness	0.10	0.15	-0.19	0.40	.442
Concrete benefit	0.16	0.18	-0.20	0.52	.320
Equal value	0.23	0.31	-0.39	0.86	.414
Particular near person	0.35	0.29	-0.23	0.92	.189
Children	0.41	0.14	0.13	0.68	.002**
Evolution	0.42	0.26	-0.09	0.93	.075

Note. Results of linear model predicting donation amounts with all low-frequency experimenter-rated argument features and the two significant high-frequency features ("Children" and "Counterarguments") as fixed effects. P values are based on likelihood ratio tests, as specified in the preregistration.

*** $p < .001$, ** $p < .01$.

Figure 2

Results of Two Linear Models Predicting Donation Amounts



Note. Results of two linear models predicting donation amounts with experimenter-rated argument features as fixed effects. Error bars represent 95% confidence intervals.

As might be expected, the distribution of features across arguments shows evidence of systematic structure: Certain argument features often go along with others, forming conceptually related clusters, in our set of 90 arguments. Of course, this raises the possibility that collinearity among the set of features would obscure true causal effects of argument features on charitable giving in the regression models presented above.

To explore this possibility further, we performed an exploratory factor analysis of all 20 experimenter-rated argument features, using the principal axis method. A Cattell scree test supported a 10-factor solution, which resulted in a few multi-item factors and several single-item factors. However, we determined that imposing a 3-factor solution generated factors that were more interpretable. This

analysis suggested that the “children” feature, which significantly predicted higher mean donations, is part of a larger cluster focused on drawing attention to particular people or causes in need (which we will refer to as “specific impact”). We refer to the second factor as “marginal utility”, as it focuses on the very favorable tradeoffs in utility available via charitable giving (i.e. a small amount of money can have a much larger impact for someone in extreme poverty than for a wealthy donor). Finally, we refer to the third factor as “equivalence of persons”, as it encompasses the notion that people’s life circumstances, such as being born into extreme poverty, are determined by chance, and we could just as easily be in another’s place (and should act accordingly, valuing other lives equally to our own).

Table 5

Results from a Factor Analysis of Experimenter-rated Features

Feature	Factor loading		
	1	2	3
Factor 1: Specificity of Impact			
Concrete benefit	**.81**	.11	-.26
Children	**.65**	.17	.05
Particular hardships	**.54**	**.31**	.03
Single hardship	**.61**	-.18	-.18
Self-interest	-.28	.00	-.13
Particular near person	.25	-.07	.13
Factor 2: Marginal Utility			
Diminishing margins for rich	.05	**.83**	.00
Massive margins for poor	.25	**.61**	-.11
Participant resources	.14	**.61**	.19
Experts	-.04	**.31**	-.20
External sources	-.02	.26	-.27
Better than nothing	.08	.23	.01
Behavioral contagion	-.07	.22	.14
Religion	-.10	.21	-.12
Common	-.02	.20	-.03
Evolution	.14	.17	-.03
Factor 3: Equivalence of Persons			
VOI, perspective taking	.12	-.04	**.66**
Luck/randomness	-.08	.01	**.42**
Equal value	.15	.15	**.32**
Counterarguments	.07	.09	-.15

Note. Factor analysis of experimenter-rated features. Factor loadings above .30 are in bold. Features loading less than .30 onto any factor are listed after the factor onto which they load highest. Note that "Particular hardships" loads onto both "Specificity of impact" and "Marginal utility".

When we ran a linear mixed-effects regression predicting donations with the three factors as fixed-effect predictors (and argument index as a random effect), we found that both the "specificity of impact" (b = 0.19, SE = 0.05, p < .001) and "marginal utility" (b = 0.18, SE = 0.08, p < .001) factors predicted increased donations. These two main clusters seemed to capture most individual features that were predictive of donations but likely obscured by collinearity in our preregistered analyses.

Table 6

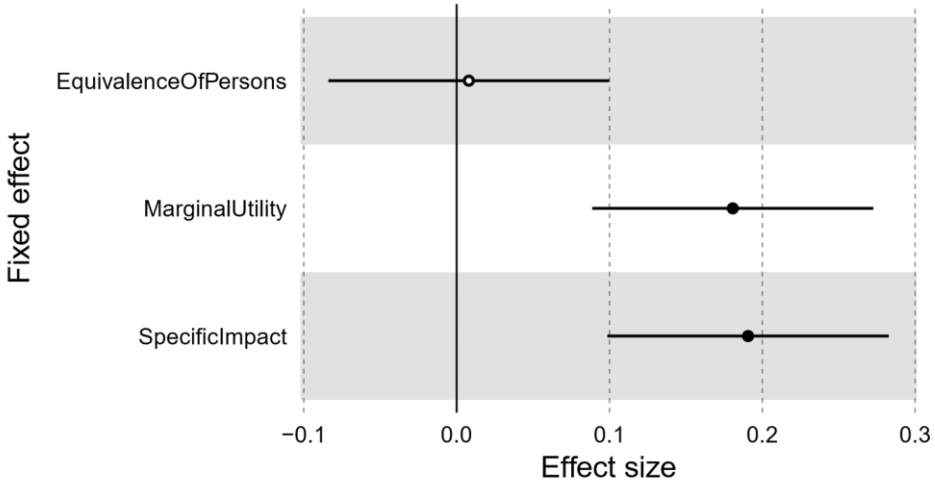
Model Summary of Linear Model Predicting Donation Amounts

Effect	Estimate	*SE*	95% CI		*p*
			LL	*UL*	
(Intercept)	3.41	0.05	3.32	3.51	<.001***
EquivalenceOfPersons	0.01	0.05	-0.09	0.10	.860
MarginalUtility	0.18	0.05	0.09	0.27	<.001***
SpecificImpact	0.19	0.05	0.10	0.28	<.001***

Note. Results of linear model predicting donation amounts with experimenter-rated argument feature factors as fixed effects. *** p < .001.

Figure 3

Results of Linear Model Predicting Donation Amounts



Note. Results of linear model predicting donation amounts with experimenter-rated argument feature factors as fixed effects. Error bars represent 95% confidence intervals.

In summary, the philosophical arguments that were most effective in persuading people to donate emphasized that a small sacrifice on the donor's part can generate a large benefit for the recipient, and also highlighted examples of specific groups of people and/or causes (e.g., trachoma-induced blindness) that stand to benefit from donations.

2.2.1.3 Participant-rated Argument Features. Next, we explored how donation behavior was related to participant ratings, drawn from each participant's responses to 16 additional questions concerning their perceptions of the argument they read. Again, as outlined in our preregistration, we ran a linear mixed-effects model to predict participants' donation amounts with all 16 participant-rated features as fixed effects and argument index as a random effect. A majority of the participant-rated features (all but five) were significantly predictive of donation to some degree.

Table 7

Model Summary of Linear Model Predicting Donation Amounts

Effect	Estimate	*SE*	95% CI		*p*
			LL	*UL*	
(Intercept)	0.75	0.18	0.39	1.11	<.001***
Hopeful	-0.09	0.03	-0.14	-0.04	<.001***
Sound	-0.09	0.03	-0.15	-0.03	.004**
Guilty	-0.09	0.02	-0.13	-0.05	<.001***
Vivid	-0.08	0.02	-0.12	-0.04	<.001***
Accusatory	-0.06	0.02	-0.10	-0.02	.005**
Easy	-0.04	0.03	-0.09	0.02	.200
Learned	-0.03	0.02	-0.06	0.01	.208
Well-written	-0.01	0.03	-0.06	0.05	.868
Kept attention	0.00	0.03	-0.05	0.05	.990
Shared	0.05	0.03	-0.00	0.11	.059
Emotional	0.08	0.03	0.02	0.13	.007**
Will think	0.08	0.02	0.04	0.13	<.001***
Agreed	0.09	0.03	0.03	0.15	.005**
Convincing	0.09	0.03	0.03	0.15	.004**
Relevant	0.24	0.03	0.19	0.29	<.001***
Made a difference	0.44	0.02	0.40	0.47	<.001***

Note. Results of linear model predicting donation amounts with all participant-rated argument features as fixed effects. *** $p < .001$, ** $p < .01$.

The directionality of some of the results might seem surprising (e.g., “hopeful”, “sound”, and “vivid” predicting lower donation in the model). However, inspection of the correlation matrix of the participant-rated features revealed high correlations among many of these features, suggesting that the model might be overfitting to collinear features. Accordingly, we again performed an exploratory factor analysis of all 16 participant-rated argument features, using the principal axis method. A Cattell scree test supported a 5-factor solution; however, we determined that imposing a 4-factor solution generated factors that were more interpretable.

Table 8

Results from a Factor Analysis of Participant-rated Features

Feature	Factor loading			
	1	2	3	4
Factor 1: Compelling				
Agreed	**.80**	.22	.23	-.07
Sound	**.75**	.25	.28	-.06
Convincing	**.70**	**.42**	.27	.00
Shared	**.60**	**.46**	**.30**	-.03
Relevant	**.58**	**.34**	.22	-.00
Hopeful	**.52**	**.47**	.29	-.10
Emotional	**.49**	**.46**	.21	.08
Factor 2: Changed Mind				
Will think	**.30**	**.68**	.23	.07
Learned	.19	**.63**	.03	.08
Made a difference	**.33**	**.47**	.08	.21
Factor 3: Engaging				
Easy	.20	-.04	**.82**	-.10
Well-written	**.37**	.21	**.70**	-.09
Kept attention	.20	.27	**.67**	-.09
Vivid	.19	**.46**	**.48**	.05
Factor 4: Brow-beating				
Guilty	.06	.19	-.08	**.74**
Accusatory	**-.42**	-.04	-.15	**.48**

Note. Factor analysis of participant-rated features. Factor loadings above .30 are in bold. Note that several items load above .3 onto more than one factor.

When all four factors were submitted to a linear mixed-effects model, three factors, which we refer to as “compelling” (b = 0.71, SE =0.03, $p < .001$), “changed mind” (b = 0.68, SE = 0.03, $p < .001$),

and "engaging" (b = 0.09, SE = 0.03, p = .003), significantly predicted increased donations. The remaining factor, "brow-beating", was not significantly predictive of donations.

Table 9

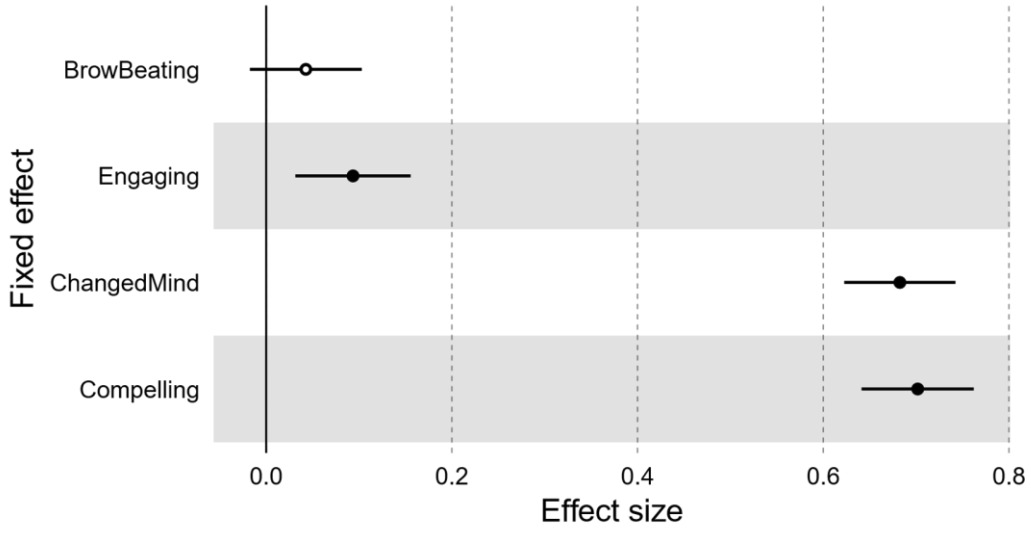
Model Summary of Linear Model Predicting Donation Amounts

Effect	Estimate	*SE*	95% CI		*p*
			LL	*UL*	
(Intercept)	3.41	0.05	3.30	3.51	<.001***
BrowBeating	0.04	0.03	-0.02	0.10	.165
Engaging	0.09	0.03	0.03	0.16	.003**
ChangedMind	0.68	0.03	0.62	0.74	<.001***
Compelling	0.70	0.03	0.64	0.76	<.001***

Note. Results of linear model predicting donation amounts with all participant-rated feature factors as fixed effects. *** $p < .001$, ** $p < .01$.

Figure 4

Results of Linear Model Predicting Donation Amounts



Note. Results of linear model predicting donation amounts with all participant-rated feature factors as fixed effects. Error bars represent 95% confidence intervals.

Taken together, these results suggest that participants may have accurately reported whether the arguments were subjectively compelling (e.g., had sound reasoning, affected them emotionally), changed their minds (e.g., taught them something new, made a difference in how much they chose to donate), and were for them stylistically engaging (e.g., ease of understanding, vividness), at least insofar as we can infer from the fact that their patterns of donation matched what one would expect given these perceptions. However, their evaluations of the "brow-beating" nature of the arguments (e.g., guilt-provoking, accusatory) were not associated with their donations to the same extent.

2.2.2 Charitable Attitudes

We now turn to our analysis of expressed charitable attitudes. As a reminder, this measure averaged the extent to which each participant agreed with two statements: 1) charitable donation is morally good, and 2) people such as themselves should donate to charity. Our analyses of this measure, all preregistered unless explicitly noted as exploratory, are discussed below.

2.2.2.1 Experimenter-rated Argument Features. First, we ran a linear mixed-effects model to predict participants' charitable attitudes with high-frequency features as fixed effects and argument index as a random effect. We found that only one feature was significantly predictive of charitable attitudes: mentioning how charitable donations offer massive margins of return for the poor, which predicted greater agreement with the charitable attitudes measure (b = 0.08, SE = 0.03, p = .013).

Table 10

Model Summary of Linear Model Predicting Charitable Attitudes

Effect	Estimate	*SE*	95% CI		*p*
			LL	*UL*	
(Intercept)	1.72	0.02	1.67	1.76	<.001***
Counterarguments	-0.01	0.03	-0.07	0.05	.761
Diminishing margins for rich	-0.00	0.04	-0.08	0.07	.908
Particular hardships	0.00	0.03	-0.06	0.06	.944
Participant resources	0.01	0.03	-0.05	0.07	.798
Experts	0.03	0.03	-0.03	0.08	.266
Children	0.04	0.03	-0.02	0.10	.192
Self-interest	0.05	0.03	-0.01	0.10	.075
Massive margins for poor	0.08	0.03	0.01	0.15	.013*

Note. Results of linear model predicting charitable attitudes with high-frequency experimenter-rated argument features as fixed effects. *** $p < .001$, * $p < .05$.

Next, we ran another linear mixed-effects model predicting participants' charitable attitudes. We retained the sole significant high-frequency feature as a fixed effect, and additionally included all 12 of the low-frequency features; the model was otherwise identical to the first. We found that two low-frequency features were predictive of charitable attitudes: a focus on one single cause/hardship (b = -0.15, SE = 0.05, p = .003), and highlighting a specific, quantified benefit of donating (e.g. three mosquito nets) (b = 0.09, SE = 0.04, p = .012). The previously significant high-frequency feature, massive margins of return, rose above the significance threshold in this model (b = -0.05, SE = 0.03, p = .060).

Table 11

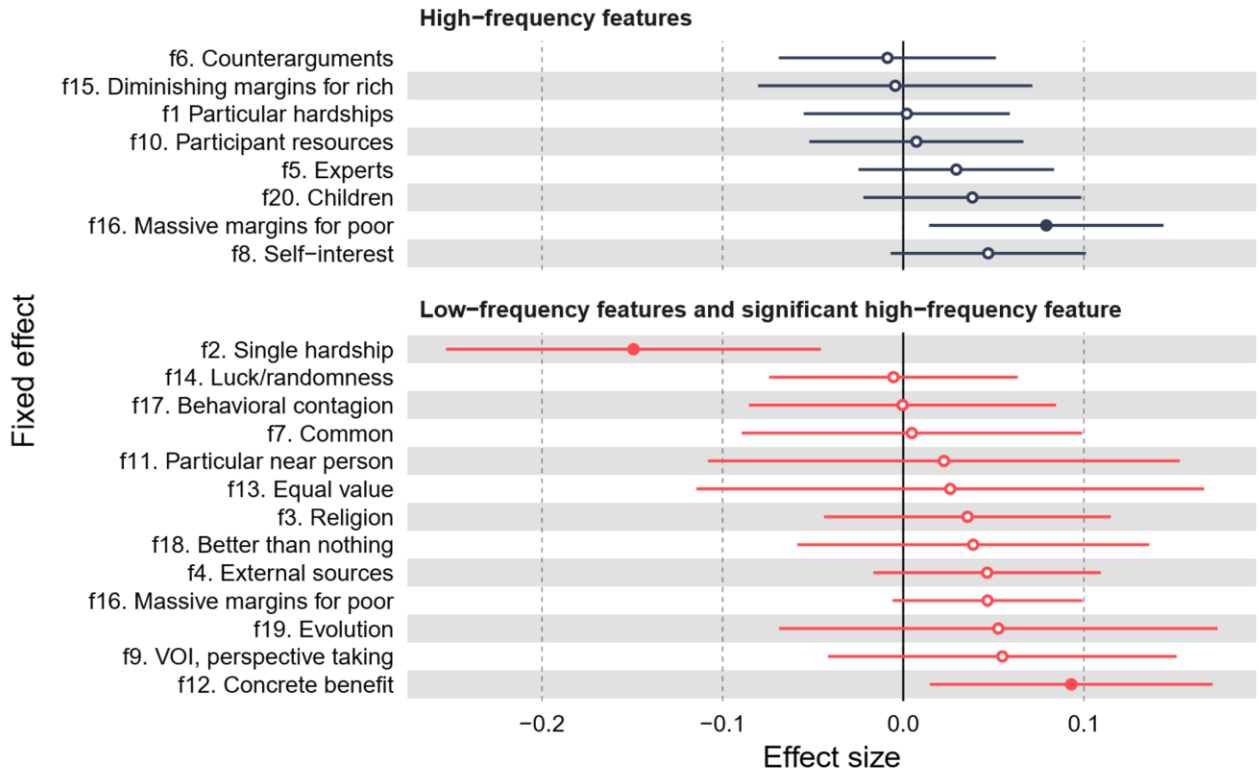
Model Summary of Linear Model Predicting Charitable Attitudes

Effect	Estimate	*SE*	95% CI		*p*
			LL	*UL*	
(Intercept)	1.74	0.02	1.70	1.77	<.001***
Single hardship	-0.15	0.05	-0.26	-0.04	.003 **
Luck/randomness	-0.01	0.04	-0.08	0.07	.866
Behavioral contagion	0.00	0.04	-0.09	0.09	.987
Common	0.01	0.05	-0.09	0.10	.911
Particular near person	0.02	0.07	-0.11	0.16	.717
Equal value	0.03	0.07	-0.12	0.17	.694
Religion	0.04	0.04	-0.05	0.12	.337
Better than nothing	0.03	0.05	-0.06	0.14	.400
External sources	0.05	0.03	-0.02	0.11	.120
Massive margins for poor	0.05	0.03	-0.01	0.10	.060
Evolution	0.05	0.06	-0.07	0.18	.355
VOI, perspective taking	0.06	0.05	-0.04	0.15	.223
Concrete benefit	0.09	0.04	0.01	0.17	.012 *

Note. Results of linear model predicting charitable attitudes with all low-frequency experimenter-rated argument features and the single significant high-frequency feature ("Massive margins for poor") as fixed effects. *** $p < .001$, ** $p < .01$, * $p < .05$.

Figure 5

Results of Two Linear Models Predicting Charitable Attitudes



Note. Results of two linear models predicting charitable giving attitudes with experimenter-rated argument features as fixed effects. Error bars represent 95% confidence intervals.

Overall, the experimenter-rated features of the philosophical arguments seemed to better predict charitable behavior (donation) than charitable attitudes. This might in part be because attitude responses were near ceiling (Good $M = 2.15$, $SD = 0.94$; Like Me $M = 1.42$, $SD = 1.33$) while donation amounts were nearer midpoint.

2.2.2.2 Participant-rated Argument Features. Finally, we ran a third linear mixed-effects model to predict participants' charitable attitudes with all 16 participant-rated features as fixed effects and argument index as a random effect. A majority of the participant-rated features were significantly predictive of attitudes to some degree.

Many of the features that one would expect to predict more positive attitudes towards charitable giving, such as agreeing with the text, finding it convincing and attention-grabbing, finding it emotionally moving and relevant to one's beliefs or goals, and agreeing that it led them to donate more of their bonus that they otherwise would have, were indeed positively associated with charitable attitudes. Meanwhile, participants' rating of a text as "accusatory" was the strongest predictor of negative attitudes, suggesting a backlash effect. Perhaps surprisingly, participants' assessment of the quality of the argument (that it had sound reasoning and was well-written) were both negatively predictive of charitable attitudes. This might again be due to the collinearity of the predictors in the model: Simple correlations between these two features and attitudes were positive, as expected (sound: $r = .32$, $p < .001$; well-written: $r = .21$, $p < .001$, not preregistered).

Table 12

Model Summary of Linear Model Predicting Charitable Attitudes

Effect	Estimate	*SE*	95% CI		*p*
			LL	*UL*	
(Intercept)	0.25	0.06	0.14	0.37	<.001***
Accusatory	-0.03	0.01	-0.05	-0.02	<.001***
Well-written	-0.03	0.01	-0.05	-0.01	.001**
Sound	-0.03	0.01	-0.04	-0.01	.014*
Learned	-0.01	0.01	-0.02	0.00	.099
Vivid	-0.01	0.01	-0.02	0.00	.179
Guilty	-0.01	0.01	-0.02	0.00	.163
Shared	-0.01	0.01	-0.03	0.01	.468
Hopeful	-0.01	0.01	-0.02	0.01	.455
Easy	-0.00	0.01	-0.02	0.02	.833
Will think	0.01	0.01	-0.01	0.02	.286
Kept attention	0.03	0.01	0.01	0.05	<.001***
Made a difference	0.04	0.01	0.03	0.05	<.001***
Convincing	0.05	0.01	0.03	0.07	<.001***
Emotional	0.05	0.01	0.03	0.07	<.001***
Agreed	0.10	0.01	0.08	0.12	<.001***
Relevant	0.16	0.01	0.14	0.18	<.001***

Note. Results of linear model predicting charitable attitudes with all participant-rated argument features as fixed effects. *** $p < .001$, ** $p < .01$, * $p < .05$.

3. General Discussion

We find that some philosophical arguments can be effective in increasing charitable giving in the context of our study. In Study 1, we identified 5 arguments as especially likely candidates to influence charitable giving, and we tested these against a neutral baseline. The average effect was to boost charitable donation from a lotteried $10 windfall by about 40%—from a baseline around $2.60 to an average gift of about $3.60. While prior research provides some reasons to be skeptical whether exposure to philosophical arguments generally improves moral behavior (Schwitzgebel & Rust, 2016), our findings

join several recent others in showing that such effects are at least possible (Buckland et al., 2022; Lindauer et al., 2020; Schwitzgebel et al., 2024).

In Study 2, we explored all 90 philosophical arguments entered into our contest, finding substantial variability in effectiveness. We attempted to explain this variability by modeling the effect of specific argument features on charitable giving. This analysis identified two clusters of argument features that commonly co-occurred and reliably boosted charitable giving. One of these clusters ("specificity of impact") included features drawing attention to specific details of the charity's influence on welfare: The hardships of the target population, the fact that children are among this population, and the specific benefits they would obtain through charitable gifts. The second of these clusters ("marginal utility") included features drawing attention to the fact that a small sacrifice on the part of the participant could generate a large benefit for those receiving the benefits of the charity.

The efficacy of the first cluster, "specificity of impact", comports with a large literature showing how concrete, personal details (including about "identifiable victims") are especially likely to motivate charitable behaviors (Jenni & Loewenstein, 1997; Kogut & Ritov, 2005; Small & Loewenstein, 2003). The efficacy of the second cluster shows that relatively abstract cost-benefit reasoning can also be persuasive. Combined, these two clusters of features suggest that consequentialist, welfare-based reasoning can effectively shift behavior, especially when supported with concrete supporting details.

Many other features were represented in our set of 90 arguments but did not show statistically reliable effects on charitable giving. Although we must interpret these null results with substantial caution, identifying several features that showed no apparent association with charitable giving may help to guide future research. A third cluster of co-occurring argument features focused on the equivalence of persons, which might be interpreted as a more "rights"-based approach than the relatively consequentialist arguments described above. These showed no evidence of motivating charitable giving in our study. We also find no evidence that it was helpful to frame arguments for charitable giving in terms of self interest (e.g., that one will "feel better" after having acted charitably).

In order to best interpret these results, it helps to contextualize them in several ways. If exposure to a certain kind of philosophical argument increases charitable giving, it is natural to conclude that people find this kind of philosophical argument relevant and at least somewhat convincing. Yet, if exposure to another kind of philosophical argument fails to increase charitable giving, the appropriate conclusion is less certain. It may be that people do not find this kind of philosophical argument relevant and convincing. However, it may instead be that people had already considered that argument on their own—in other words, that its influence was "baked in" to their default responses in the control condition. It may also be that the specific expression of the argument in our stimuli was unpersuasive, while a different expression of it would be more persuasive. Furthermore, there are several important limitations to our design: Our participant population comprises exclusively US residents, and likely skews politically liberal (Levay et al., 2016); MTurk participants might differ from the general population in several respects (Krupnikov & Levine, 2014; McCredie & Morey, 2019), and specifically their willingness to engage in online tasks for small amounts of money might relate to their willingness or unwillingness to donate similarly small amounts of money to charity; opportunities to give were presented immediately after exposure to philosophical arguments, and so may not endure; we asked people to make small donations of a lottery-based windfall profit, which has few real-world analogues; and, participants were aware that they were participating in a study and were therefore subject to task demands.

In addition to offering participants an opportunity to give a portion of a lotteried $10 windfall to charity, we also assessed their attitudes towards charitable giving. We consistently found that people have very positive attitudes towards charitable giving, describing it as "morally good" and agreeing that "people like me" should give to charitable causes. Yet, the effects of exposure to arguments were smaller and less consistent for these measures, perhaps in part because of the compressed range for causal effects given that attitudes were high at baseline.

For several decades, psychologists have debated whether, and when, reasoned argument can change moral thought and behavior (May, 2018; Paxton & Greene, 2010). And, for several millennia, philosophers have honed reasoned arguments about morality, including some designed to change moral

thought and behavior. We show that some philosophical arguments can have a substantial and immediate effect on the amount of money that US participants give to charity from a small, lotteried windfall profit. We also identify specific features of arguments that reliably drive such effects, which emphasize in both concrete and abstract ways a simple, powerful idea: With charity, one can generate large and specific benefit to others at a small cost to oneself.

Open Practices

The research question, methods, and analysis plan for Study 2 were preregistered (https://aspredicted.org/9mtf-tcmh.pdf) on 2022-05-31, prior to data collection which began on 2022-06-10. All study materials, data, and analysis scripts for both Study 1 and Study 2 are publicly available (https://osf.io/ujzfc/files/osfstorage?view_only=aa1747631e7b4350b8c5203ad6655642).

References

Bagozzi, R. P., & Moore, D. J. (1994). Public Service Advertisements: Emotions and Empathy Guide Prosocial Behavior. *Journal of Marketing*, *58*(1), 56–70. https://doi.org/10.2307/1252251

Buckland, L., Lindauer, M., Rodríguez-Arias, D., & Véliz, C. (2022). Testing the Motivational Strength of Positive and Negative Duty Arguments Regarding Global Poverty. *Review of Philosophy and Psychology*, *13*(3), 699–717. https://doi.org/10.1007/s13164-021-00555-4

Cacioppo, J. T., Petty, R. E., & Feng Kao, C. (1984). The Efficient Assessment of Need for Cognition. *Journal of Personality Assessment*, *48*(3), 306–307. https://doi.org/10.1207/s15327752jpa4803_13

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*(4), 444–474. https://doi.org/10.1080/10495142.2018.1452828

Grodeck, B., & Schoenegger, P. (2023). Demanding the morally demanding: Experimental evidence on the effects of moral arguments and moral demandingness on charitable giving. *Journal of Behavioral and Experimental Economics*, *103*, 101988. https://doi.org/10.1016/j.socec.2023.101988

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*(4), 814–834. https://doi.org/10.1037/0033-295X.108.4.814

Hou, T., Ding, X., & Yu, F. (2024). The moral behavior of ethics professors: A replication-extension in Chinese mainland. *Philosophical Psychology*, *37*(2), 396–427. https://doi.org/10.1080/09515089.2022.2084057

Jalil, A. J., Tasoff, J., & Bustamante, A. V. (2020). Eating to save the planet: Evidence from a randomized controlled trial using individual-level food purchase data. *Food Policy*, *95*, 101950. https://doi.org/10.1016/j.foodpol.2020.101950

Jalil, A. J., Tasoff, J., & Bustamante, A. V. (2023). Low-cost climate-change informational intervention reduces meat consumption among students for 3 years. *Nature Food*, *4*(3), 218–222.

https://doi.org/10.1038/s43016-023-00712-1

Jenni, K. E., & Loewenstein, G. (1997). Explaining the Identifiable Victim Effect. *Journal of Risk and Uncertainty*, *14*, 235–257.

Kogut, T., & Ritov, I. (2005). The "identified victim" effect: An identified group, or just a single individual? *Journal of Behavioral Decision Making*, *18*(3), 157–167. https://doi.org/10.1002/bdm.492

Kogut, T., & Ritov, I. (2011). The identifiable victim effect: Causes and boundary conditions. In D. M. Oppenheimer & C. Y. Olivola, *Experimental Approaches to the Study of Charity* (pp. 133–145). Psychology Press.

Krupnikov, Y., & Levine, A. S. (2014). Cross-Sample Comparisons and External Validity. *Journal of Experimental Political Science*, *1*(1), 59–80. https://doi.org/10.1017/xps.2014.7

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Levay, K. E., Freese, J., & Druckman, J. N. (2016). The Demographic and Political Composition of Mechanical Turk Samples. *Sage Open*, *6*(1), 2158244016636433. https://doi.org/10.1177/2158244016636433

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*(3), 413–420. https://doi.org/10.1017/S1930297500007208

May, J. (2018). *Regard for Reason in the Moral Mind*. Oxford University Press.

McCredie, M. N., & Morey, L. C. (2019). Who Are the Turkers? A Characterization of MTurk Workers Using the Personality Assessment Inventory. *Assessment*, *26*(5), 759–766. https://doi.org/10.1177/1073191118760709

Mercier, H. (2011). What good is moral reasoning? *Mind & Society*, *10*(2), 131–148. https://doi.org/10.1007/s11299-011-0085-6

Nichols, S., & Knobe, J. (2007). Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions. *Noûs*, *41*(4), 663–685.

Paxton, J. M., & Greene, J. D. (2010). Moral Reasoning: Hints and Allegations. *Topics in Cognitive Science*, *2*(3), 511–527. https://doi.org/10.1111/j.1756-8765.2010.01096.x

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*, 39–50. https://doi.org/10.1016/j.cognition.2018.06.011

R Core Team. (2023). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Schönegger, P., & Wagner, J. (2019). The moral behavior of ethics professors: A replication-extension in German-speaking countries. *Philosophical Psychology*, *32*(4), 532–559. https://doi.org/10.1080/09515089.2019.1587912

Schwitzgebel, E. (2019). *A Theory of Jerks and Other Philosophical Misadventures*. MIT Press.

Schwitzgebel, E., Cokelet, B., & Singer, P. (2020). Do ethics classes influence student behavior? Case study: Teaching the ethics of eating meat. *Cognition*, *203*, 104397. https://doi.org/10.1016/j.cognition.2020.104397

Schwitzgebel, E., Cokelet, B., & Singer, P. (2023). Students Eat Less Meat After Studying Meat Ethics. *Review of Philosophy and Psychology*, *14*(1), 113–138. https://doi.org/10.1007/s13164-021-00583-0

Schwitzgebel, E., McVey, C., & May, J. (2022). Engaging charitable giving: The motivational force of narrative versus philosophical argument. *Philosophical Psychology*, *0*(0), 1–36. https://doi.org/10.1080/09515089.2022.2088340

Schwitzgebel, E., McVey, C., & May, J. (2024). Engaging charitable giving: The motivational force of narrative versus philosophical argument. *Philosophical Psychology*, *37*(5), 1240–1275. https://doi.org/10.1080/09515089.2022.2088340

Schwitzgebel, E., & Rust, J. (2016). The Behavior of Ethicists. In *A Companion to Experimental*

Philosophy (pp. 225–233). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118661666.ch15

Small, D. A., & Loewenstein, G. (2003). Helping a Victim or Helping the Victim: Altruism and Identifiability. *Journal of Risk and Uncertainty*, *26*(1), 5–16. https://doi.org/10.1023/A:1022299422219